Structurally Engineered Cytochromes *c* with Novel Ligand-Binding Properties

Thesis by

Kara Lynne Bren

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1996

(Submitted November 14, 1995)

Acknowledgment

My experience at Caltech has been greatly enriched by the interactions I have had with many different people. The support of my friends and family and the collaborations I have enjoyed with other scientists helped make my graduate years memorable and productive.

First and foremost, I would like to thank Harry Gray for his confidence in me, through which I gained more confidence in myself. I also thank him for teaching me by sharing his insights and enthusiasm for science. I especially appreciate the support and kindness he extended to me during my time at Caltech.

I was fortunate to work with many bright people in Florence. I thank Ivano Bertini, without whose optimism we would not have met a great scientific challenge. I also thank Thep for his valuable contribution to the project and for his hospitality during my visit. I am grateful to the many members of the Bertini group who made me feel welcome and helped introduce me to NMR. In particular, I am glad that I had the chance to work with Paola, who is not only a great NMR spectroscopist but also a great friend.

I am grateful to Yi Lu and Danny Casimiro for inviting me to work with them on the yeast expression system. I appreciate the many patient hours Yi spent to teach me about site-directed mutagenesis. I also acknowledge Jack Richards and the people in the Richards labs for their help in this part of the work.

When I began the MCD collaboration with John Dawson, little did I know that I would be fortunate enough to work with someone like Bart Hawkins. I thank him for his scientific insights, but what I appreciate the most is the friendship between us that developed.

I thank the many members of the Gray group who have helped me with my research. In particular, I thank Debbie Wuttke for getting me started in lab by introducing me to semisynthesis. I also was fortunate to work with two talented undergraduates, Natalie Austin and Grace Yang, as part of the SURF program.

Many people not named above have extended personal support to me during my years as a graduate student. I cannot name you all, but I would like to single out a few. Don has been a wonderful friend, and I know that we will continue to be friends for a long time. I have enjoyed the long chats I've had with Erik, and also I am grateful to him for being my "car guy." I thank Gautam, Regan, Catherine, and Michele for their continued correspondence and for sharing some wonderful times in new places. I thank Adrian for his kindness, support, and friendship--I am always looking forward to our next trip together!

Finally, I reserve a special thanks for my family. Even from 2000 miles away, their love and support has meant a lot to me during my time here.

Abstract

Semisynthesis of horse heart cytochrome *c* and site-directed mutagenesis of *Saccharomyces cerevisiae* (*S. c.*) iso-1-cytochrome *c* have been utilized to substitute Ala for the cytochrome *c* heme axial ligand Met80 to yield ligand-binding proteins (horse heart Ala80cyt *c* and *S. c.* Ala80cyt *c*) with spectroscopic properties remarkably similar to those of myoglobin. Both species of Fe(II)Ala80cyt *c* form exceptionally stable dioxygen complexes with autoxidation rates 10–30x smaller and O_2 binding constants ~ 3x greater than those of myoglobin. The resistance of O_2-Fe(II)Ala80cyt *c* to autoxidation is attributed in part to protection of the heme site from solvent as exhibited by the exceptionally slow rate of CO binding to the heme as well as the low quantum yield of CO photodissociation.

UV/vis, EPR, and paramagnetic NMR spectroscopy indicate that at pH 7 the Fe(III)Ala80cyt *c* heme is low-spin with axial His-OH^- coordination and that below pH ~ 6.5, Fe(III)Ala80cyt *c* is high-spin with His-H_2O heme ligation. Significant differences in the pH dependence of the ^1H NMR spectra of *S. c.* Fe(III)Ala80cyt *c* compared to wild-type demonstrate that the axial ligands influence the conformational energetics of cytochrome *c*.

^1H NMR spectroscopy has been utilized to determine the solution structure of the cyanide derivative of *S. c.* Fe(III)Ala80cyt *c*. 82% of the resonances in the ^1H NMR spectrum of *S. c.* CN-Fe(III)Ala80cyt *c* have been assigned through 1D and 2D experiments. The RMSD values after restrained energy minimization of the family of 17 structures obtained from distance geometry calculations are 0.68 ± 0.11 Å for the backbone and 1.32 ± 0.14 Å for all heavy atoms. The solution structure indicates that a tyrosine in the "distal" pocket of CN-Fe(III)Ala80cyt *c* forms a hydrogen bond with the Fe(III)-CN unit, suggesting that it may play a role analogous to that of the distal histidine in myoglobin in stabilizing the dioxygen adduct.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Abbreviations

1D	one-dimensional
2D	two-dimensional
atm	atmosphere
CcP	cytochrome *c* peroxidase
CD	circular dichroism
COSY	correlation spectroscopy
cyt *c*	cytochrome *c*
DG	distance geometry
DNA	deoxyribonucleic acid
DSS	sodium 2,2-dimethyl-2-silapentane-5-sulfonate
EDTA	ethylenediaminetetraacetic acid
EPR	electron paramagnetic resonance
FID	free induction decay
FPLC	fast protein liquid chromatography
h	hour
Hb	hemoglobin
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
HOMO	highest occupied molecular orbital
H. S.	high-spin
IDA	iminodiacetic acid
kb	kilo-basepair
kDa	kilodalton
L. S.	low-spin
Mb	myoglobin
MCOSY	magnitude COSY
min.	minutes
mmHg	millimeters of mercury
mV	millivolt
MW	molecular weight
NHE	normal hydrogen electrode
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	two-dimensional nuclear Overhauser effect spectroscopy
OD	optical density
PDB	protein data bank

phen	1,10-phenanthroline
P_i	phosphate
Pxs	peroxidases
REDAC	redundant angle calculation
REM	restrained energy minimization
RMD	restrained molecular dynamics
RMSD	root mean squared deviation
rpm	revolutions per minute
s	second
S. c.	*Saccharomyces cerevisiae*
TOCSY	total correlation spectroscopy
TPPI	time-proportional phase incrementation
Tris	tris-hydroxymethyl aminomethane
UV/vis	ultraviolet/visible
WEFT	water-eliminated Fourier transform
WT	wild-type

Chapter 1

Engineering Heme Protein Active Sites, an Introduction

The heme group displays remarkable versatility in nature, serving as the active site of proteins that perform electron transfer (cytochromes *c*),[1] oxygen storage and transport (myoglobins and hemoglobins),[2] and redox catalysis (peroxidases,[3] cytochrome P-450[4]). The ability of the same prosthetic group (Figure 1.1) to perform these different functions is due to the influence of the protein matrix surrounding the heme. Protein engineering has served as a valuable technique for elucidating the role of active site amino acids in heme proteins such as hemoglobin (Hb), myoglobin (Mb), peroxidases, and cytochrome *c* (cyt *c*). In addition to probing protein function, protein engineering has also been used to alter protein function. This allows the use of one protein as a model system in the study of another. With the tools of protein engineering and the expanding base of information on biochemical structure/function relationships in hand, in the future it may be possible to design and engineer proteins to exhibit activities such as specific ligand binding, small molecule scavenging, or highly stereospecific redox catalysis. The successes already met in both probing and altering the function of proteins demonstrate the utility of protein engineering for both the elucidation of structure/function relationships and the production of proteins with novel properties.

Engineering Heme Protein Active Sites

Mbs and Hbs are oxygen storage and transport proteins found in a wide variety of organisms (Figure 1.2A).[2] To function in this capacity, Hbs and Mbs must form stable heme-dioxygen adducts, which is possible only for the ferrous oxidation state of heme. Because free ferrous-oxy heme autoxidizes rapidly to the ferric state,[5] the polypeptide surrounding the heme must play a crucial role in stabilizing the ferrous-oxy form (Figure 1.3A, Figure 1.4). Before the development of site-directed mutagenesis, naturally occurring mutant proteins provided a medium for investigating the role played by particular active-site residues in modulating heme function. The Hbs M are a class of human Hbs that display significantly altered reactivity toward dioxygen compared to

Figure 1.1 The heme group consists of an iron protoporphyrin IX bound to an iron. The iron is most commonly found in the +2 (ferrous) and +3 (ferric) oxidation states.

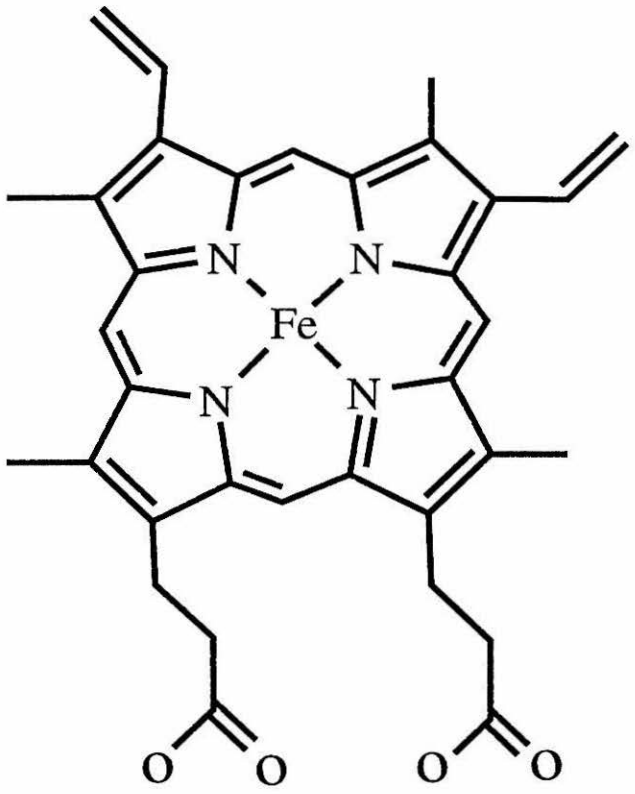
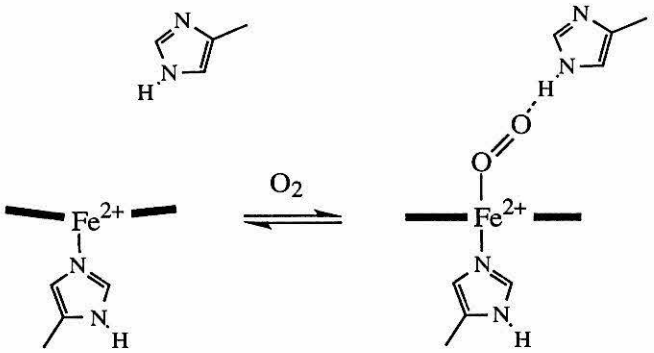

N
N
Fe
N
N
O
O
O
O

Figure 1.2 Schematic illustration of the function of three representative heme proteins. (A) Myoglobin and hemoglobin bind dioxygen reversibly and serve as dioxygen storage and transport proteins. (B) Cytochrome *c* peroxidase converts hydrogen peroxide to water and oxidizes cytochrome *c*. Peroxidases in general are capable of catalyzing the peroxide-dependent oxidation of a wide variety of organic and inorganic compounds.[3] (C) Cytochrome *c* is a vital link in the mitochondrial respiratory chain, shuttling electrons between cytochrome *c* reductase (the cytochrome bc_1 complex) and cytochrome *c* oxidase.

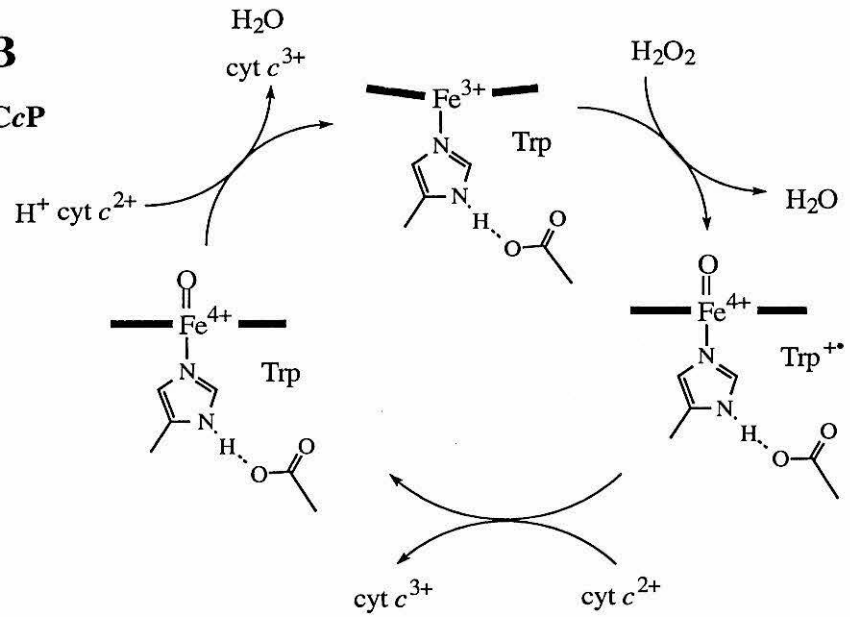
A

Mb, Hb



B

CcP



C

cyt *c*

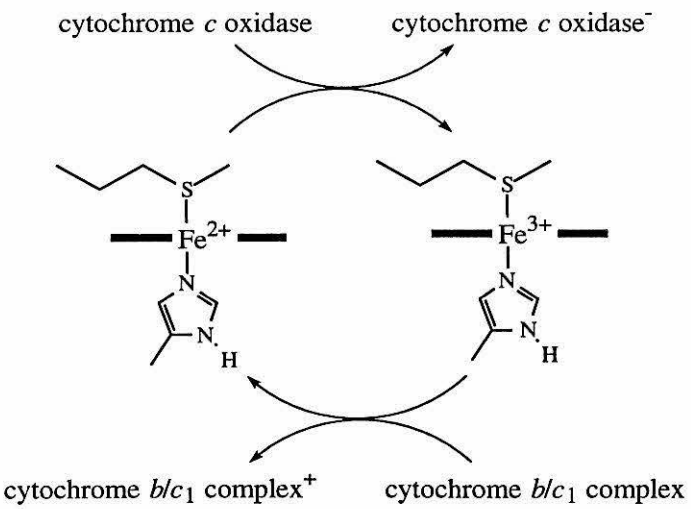


Figure 1.3 Active-site residues from the X-ray crystal structures of (A) O_2-Fe(II)sperm whale myoglobin (PDB identifier 1MBO), (B) *Saccharomyces cerevisiae* Fe(III)cytochrome *c* peroxidase (PDB identifier 2CYP), (C) *Saccharomyces cerevisiae* Fe(III)cytochrome *c* (PDB identifier 1YCC).

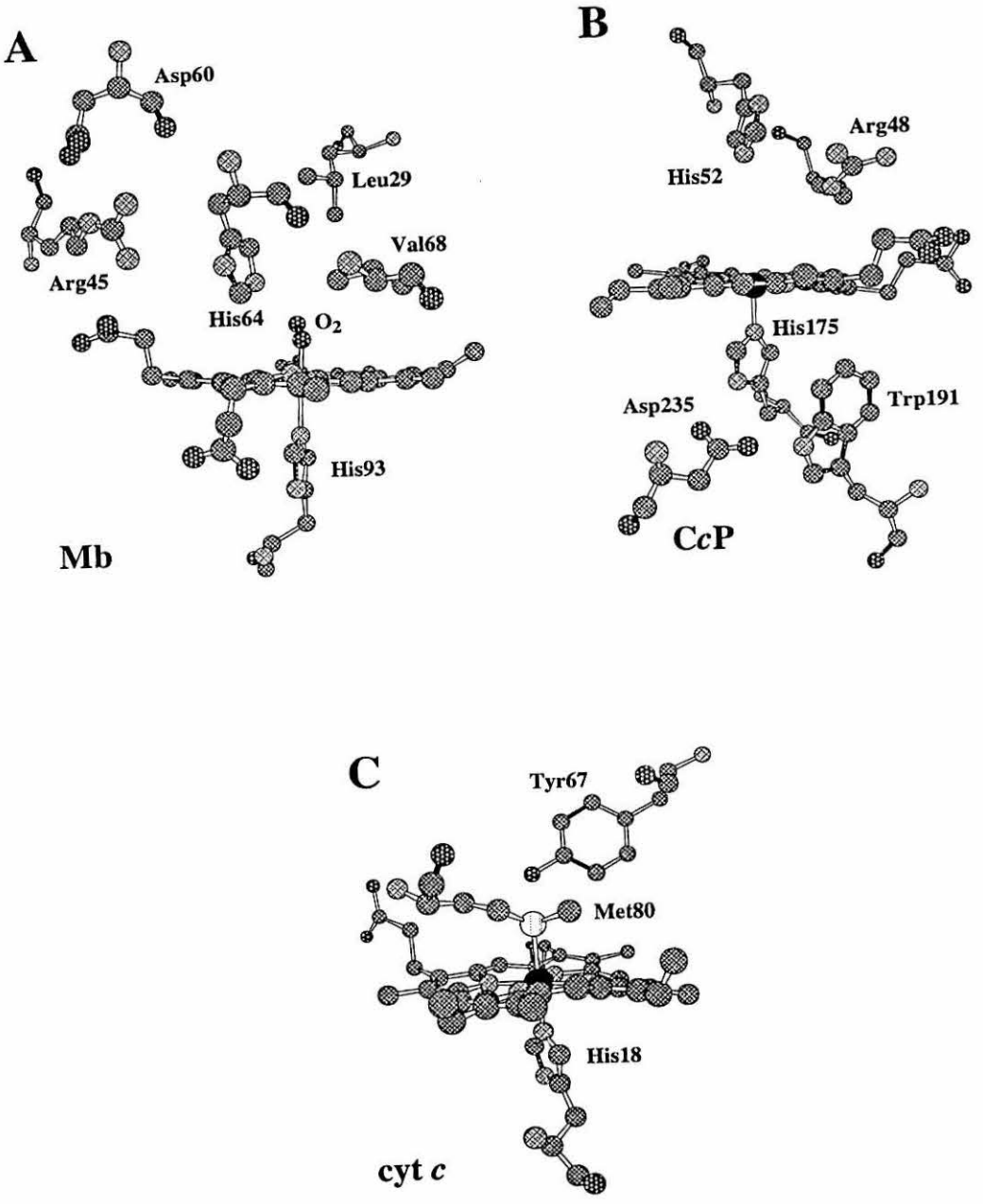
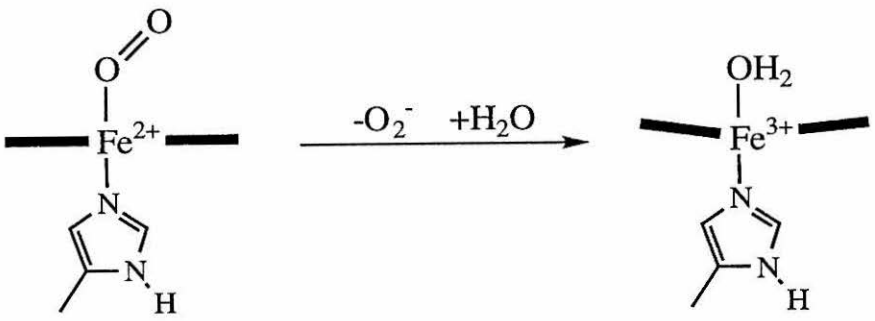

A
Asp60
Leu29
Arg45
Val68
His64
O_2
His93
Mb
B
Arg48
His52
His175
Asp235
Trp191
CcP
C
Tyr67
Met80
His18
cyt c

Figure 1.4 Illustration of the autoxidation of a ferrous-oxy heme to the ferric (Fe(III)-H_2O) form.


O
O
Fe2+
N
N
H
-O2-
+H2O
OH2
Fe3+
N
N
H

normal human Hb.[6] The four Hbs in this class have Tyr in place of His residues normally at the heme site.[7] All Hbs M show a high rate of autoxidation of those hemes that are bound to the abnormal polypeptide chains.[6] The X-ray crystal structure of one of the Hbs M in which the distal His is replaced by Tyr shows that the distal Tyr binds to the ferric iron.[8] These early studies support the hypothesis that the distal His is important for stabilizing the ferrous-oxy form of Mb and Hb.[9]

Advances in site-directed mutagenesis of Mb and Hb have led to the production of a wide variety of active-site mutants that have been used to explore how the polypeptide stabilizes ferrous-oxy Mb and Hb and discriminates against the biologically deleterious CO binding.[10,11] Mb is especially attractive for this work because its monomeric nature simplifies structure/function studies. Mutation of the Mb distal His (position E7 in helical notation[12]) reduces dioxygen affinity,[11,13,14] increases the rate of oxyheme autoxidation,[13] and increases the affinity of the heme for CO.[14] These results support the hypothesis that the distal His serves to stabilize bound O_2 through hydrogen-bonding, while inhibiting CO binding.[9] Engineering other distal pocket residues has also led to Mb mutants with altered reactivity. Most notably, Leu29Phe (position B10) sperm whale Mb has an unusually high oxygen binding constant and low autoxidation rate.[15] Leu29Phe Mb serves as a model for elephant Mb, which utilizes a B-helix Phe ring in stabilizing bound O_2.[16]

Protein engineering has afforded Mb mutants that serve as spectroscopic and functional models of other heme proteins. In particular, altering Mb heme ligation results in significant changes in heme function. To perform as a dioxygen-storage protein, the Mb heme must have a site at the heme iron capable of binding small exogenous ligands. The engineering of an amino acid that may serve as a ligand into the distal pocket of Mb has in some cases resulted in a heme iron with a coordination sphere saturated by the heme and protein-donated ligands. For example, substitution of the distal His64 with Tyr in sperm whale Mb results in a protein that mimics the abnormal Hbs M in which the

distal Tyr binds to the heme iron.[17] The function of horse heart Mb is changed from ligand-binding to electron transfer by mutating distal pocket residue Val68 into His.[18] This variant exhibits spectroscopic and electrochemical similarities to the bis-His electron transfer protein cytochrome b_5. While native Mb as well as other Mbs with altered axial ligation show oxidation state-dependent ligation, His68 directly coordinates to the heme iron in both oxidation states of the variant.[18]

A slightly different strategy in the use of Mb as a model for other heme proteins comes in replacement of the strictly conserved heme-binding proximal His93 with other amino acids, in particular with other potentially ligating side-chains. Substitution of Cys for His93 produces a Cys-ligated heme with spectroscopic properties (in the ferric state) nearly identical with those of cytochrome P-450.[19] Reaction of ferri-Cys93Mb with peroxides results in an enhanced amount of heterolytic O–O bond cleavage relative to that seen for wild-type (WT) Mb which has a proximal His,[19] thus supporting the proposal that a strong electron donor such as a Cys facilitates heterolytic O–O bond cleavage.[20] The proximal ligand mutant His93Tyr has been produced in human,[19] sperm whale,[21] and horse heart Mb.[22] Engineering Tyr into position 93 gives a protein that in the oxidized form has spectroscopic signatures very similar to those of catalase. A mutant of sperm whale Mb in which the proximal His is replaced with Gly results in a proximal-side ligand binding pocket.[23] Phenol and ethanethiol can bind at the proximal site of Gly93Mb and serve as mimics of a proximal Tyr and Cys, respectively. In addition, it has been demonstrated that other bases can bind in this pocket; this presents the opportunity to explore the effect of heme ligation by bases other than those included in the repertoire of the 20 encoded amino acids.

The core components of the Mb active site, a His-bound heme and a distal (non-ligating) His, are the same as those of cytochrome *c* peroxidase (C*c*P). However, while Mb serves as a dioxygen storage protein, C*c*P is a catalyst for the peroxide-dependent oxidation of cyt *c* (Figure 1.2B). Although the distal His in Mb is believed to help

stabilize the ferrous-oxy species, ferrous-oxy C*c*P autoxidizes rapidly, and the distal His plays a different role: it has been shown through site-directed mutagenesis to be required for the enzyme's efficient reaction with peroxide to form compound ES (a Fe(IV)=O heme and protein radical cation).[24] Through systematic mutagenesis studies, other distal amino acids (Figure 1.3B) have been demonstrated to tune C*c*P heme function. Distal residue Arg48 has been implicated in the stabilization of compound ES by acting with His52 to compensate for the charge separation produced by the heterolytic cleavage of the peroxide O–O bond.[25] Site-directed mutagenesis helped identify Trp191 as the site of the protein radical in compound ES.[26] Removal of the Trp191 indole functionality by mutation to Gly produces a heme pocket that binds small molecules, demonstrating the creation of a substrate pocket.[27]

Significant changes in C*c*P spectroscopy and reactivity result when mutations that affect the heme ligands are made. The strength of the heme iron-ligand bond has been proposed to distinguish a dioxygen-binding protein from an oxygenase or a peroxidase.[28] Although C*c*P (as well as other peroxidases) has His ligation as do Mb and Hb, systematic mutagenesis studies, combined with spectroscopic work, have demonstrated that the proximal His of C*c*P has significant imidazolate character, thus giving more "push" to aid in breaking the peroxide O–O bond.[29] Hydrogen-bonding of Asp235 to His175 is believed to be important in causing the imidazole to act as a stronger electron donor. Mutation of this Asp to Asn or Ala results in dramatic changes of C*c*P heme properties and reactivity, including causing the normally pentacoordinate heme to bind solvent (water or hydroxide) and also significantly reducing peroxidase activity.[29]

A proximal His mutant of C*c*P analogous to the His93Gly Mb variant has been reported. Replacement of the C*c*P proximal His (His175) with Gly gives a bis-aquo heme complex that has been observed to bind imidazole and N-methylimidazole at the proximal site.[30] While His175Gly C*c*P displays no peroxidase activity, imidazole binding restores activity, although not to 100%. Alteration of His175 to Gln yields a C*c*P

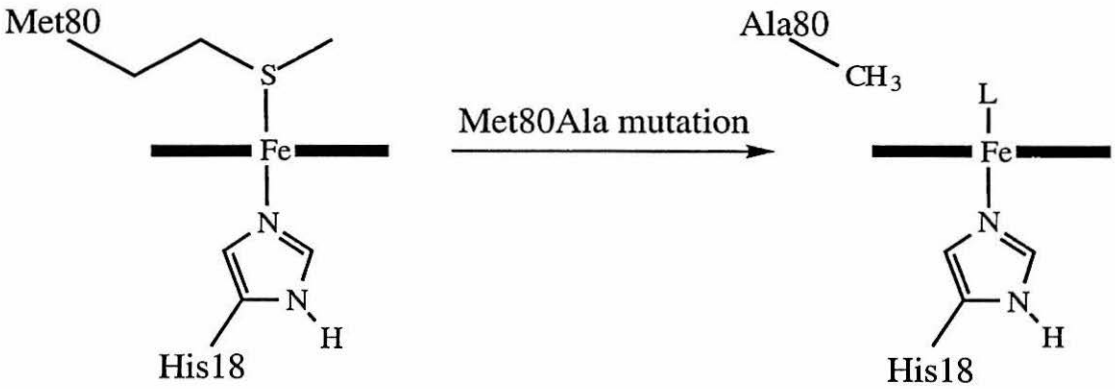
with nearly full catalytic activity, calling into question ideas on the importance of the conserved proximal His.[31] His175 has also been replaced with Glu and Cys. In the case of the Glu mutant, significant activity is retained as was observed for His175Gln. Mutation of His175 to Cys, however, leads to loss of activity. The His175Cys mutant does not have an absorption spectrum similar to cytochrome P-450. This is attributable to the oxidation of Cys175 to cysteic acid, as observed in the X-ray crystal structure of the mutant.[32]

While Mb and C*c*P function by binding small molecules (dioxygen and peroxide), the polypeptide surrounding the cyt *c* heme (Figure 1.3C) is designed to prevent small molecule binding which would interfere with its function of efficiently shuttling electrons from the cytochrome b/c_1 complex to cytochrome *c* oxidase (Figure 1.2C).[1] The heme iron in cyt *c* is coordinately saturated due to the donation of two protein axial ligands, His18 and Met80. Met80 has been shown to be important for maintaining a relatively high reduction potential (+260 mV vs. NHE) to provide the correct driving force for its biological electron transfer reactions. A number of cyt *c* mutants in which the axial Met has been altered have been reported and all have significantly lower reduction potentials than native cyt *c*.[33,34] Some of these mutants serve as mimics of other heme proteins. Substitution of the axial heme ligand Met80 with His by semisynthesis of horse heart cyt *c* results in a cytochrome b_5-like protein with bis-His ligation and a reduction potential of 41 mV vs. NHE, compared to 5 mV for native cytochrome b_5.[33,35] Semisynthetic Cys80cyt *c* has spectroscopic properties similar to the imidazole derivative of cytochrome P-450 as well as a low reduction potential of -390 mV (compared to -360 to -170 mV for cytochrome P-450).[33]

Thesis Research

The subject of this work is a position 80 cyt *c* mutant in which the iron-bound Met has been replaced by the non-ligating Ala to give Ala80cyt *c* (Figure 1.5). This mutant

Figure 1.5 Illustration of the mutation of the cytochrome *c* axial ligand Met80 to Ala. This mutation replaces the heme iron-binding thioether side chain of Met with the small, non-binding methyl group of Ala, thus opening a site at the heme iron capable of binding a variety of small exogenous ligands, L.



L = O_2, CO, NO to Fe(II)
L = H_2O, OH^-, CN^-, N_3^- to Fe(III)

has been produced both by semisynthesis of horse heart cyt *c*[34,36] and site-directed mutagenesis of *Saccharomyces cerevisiae* (*S. c.*) iso-1-cyt *c*,[37] and is capable of binding small molecules including dioxygen. Remarkably, the dioxygen adduct of Ala80cyt *c* is more stable both kinetically and thermodynamically than those reported for most mammalian Mbs. Because the polypeptide matrix surrounding the heme has been demonstrated to be instrumental in tuning the activity of hemes, we look to the residues in the heme pocket of Ala80cyt *c* for an explanation of how it is able to stabilize the Fe-O_2 unit. While the immediate active site of Ala80cyt *c* (heme and axial His) resembles that of Mb, the residues surrounding the heme are quite different. Most globins have a distal His that is needed for stabilizing bound O_2.[2] Cyt *c*, however, has no histidine in the distal pocket. Compared to the distal cavities of globins and peroxidases, that of cyt *c* is quite hydrophobic with only one residue capable of hydrogen bonding, Tyr67. This Tyr is known to hydrogen-bond with the iron-bound Met80 thioether sulfur in WTcyt *c*;[38] it is of interest to learn whether this residue can act as a "distal His" in Ala80cyt *c*.

This work has centered on the production and characterization of Ala80cyt *c* with the particular goal of elucidating the mechanism of stabilization of the Fe(II)-O_2 species. Chapters 2 and 3 discuss the preparation of Ala80cyt *c* by semisynthesis of horse heart cyt *c* (Chapter 2) and site-directed mutagenesis of yeast iso-1-cyt *c* (Chapter 3). In addition, the ligand-binding properties of these cyts *c* are reported in these chapters. Chapter 2 also includes results on a related semisynthetic mutant, Phe67Ala80cyt *c*. Chapter 4 is dedicated to the investigation of the pH-dependent behavior of Fe(III)Ala80cyt *c* compared to that exhibited by wild-type; UV/vis spectroscopy and paramagnetic ^{1}H NMR spectroscopy are used to this end. These results indicate that metal-ligand interactions can significantly influence the conformational energetics of cyt *c*.

The cyanide derivative of *S. c.* Fe(III)Ala80cyt *c* was found to be particularly amenable for a detailed ^{1}H NMR characterization, which is presented in Chapters 5 and

6. Chapter 5 focuses on the assignment and evaluation of the paramagnetically-shifted proton resonances of the heme, His18, and Tyr67 of *S. c.* CN-Fe(III)Ala80cyt *c*. Chapter 6 presents the detailed assignment of the ^{1}H NMR spectrum of *S. c.* CN-Fe(III)Ala80 and the resulting solution structure. The NMR structure determination provides important information on how Ala80cyt *c* is capable of acting as an effective dioxygen storage protein.

References and Notes

(1) Moore, G. R.; Pettigrew, G. W. *Cytochromes c; Evolutionary, Structural and Physicochemical Aspects*; Springer-Verlag: Berlin, 1990.

(2) Antonini, E.; Brunori, M. *Hemoglobin and Myoglobin in Their Reactions with Ligands*; North Holland Publishing Company: Amsterdam, 1971.

(3) Yonetani, T. In *The Enzymes*, Boyer, P. D., Ed.; Academic Press: New York, 1976; Vol. 13, pp 345-361.

(4) Ortiz de Montellano, P. R. *Cytochrome P-450: Structure, Mechanism, and Biochemistry*; Plenum Press: New York, 1986.

(5) While free heme autoxidizes to form μ-oxo dimers, autoxidation of Mb and Hb occurs through a different pathway due to the protection of the heme by the globin: Brantley, R. E., Jr.; Smerdon, S. J.; Wilkinson, A. J.; Singleton, E. W.; Olson, J. S. *J. Biol. Chem.* **1993,** *268,* 6995-7010.

(6) Ranney, H. M.; Nagel, R. L.; Heller, P.; Udem, L. *Biochim. Biophys. Acta* **1968,** *160,* 112-115. Hayashi, N.; Motokawa, Y.; Kiruchi, G. *J. Biol. Chem.* **1966,** *241,* 79-84. Suzuki, T.; Hayashi, A.; Yamamura, Y.; Enoki, Y.; Tyuma, I. *Biochem. Biophys. Res. Commun.* **1965,** *19,* 691-695. Suzuki, T.; Hayashi, A.; Shimizu, A.; Yamamura, Y. *Biochim. Biophys. Acta* **1966,** *127,* 280-282.

(7) The His residues replaced in the Hbs M are known as "distal" and "proximal." The distal His is located on the side of the heme to which O_2 (as well as other ligands) binds. This side of the heme is also referred to as the distal side. The proximal His is bound to the iron on the other (proximal) side of the heme.

(8) Pulisinelli, P. D.; Perutz, M. F.; Nagel, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1973,** *70,* 3870-3874.

(9) Pauling, L. *Nature* **1964,** *203,* 182-183. Collman, J. P. *Acc. Chem. Res.* **1977,** *10,* 265-273. Phillips, S. E. V. *J. Mol. Biol.* **1980,** *142,* 531-554.

(10) Nagai, K.; Luisi, B.; Shih, D.; Miyazaki, G.; Imai, K.; Poyart, C.; De Young, A.; Kwiatkowsky, L.; Noble, R. W.; Lin, S.-H.; Yu, N.-T. *Nature* **1987,** 329, 858-860.

(11) Olson, J. S.; Mathews, A. J.; Rohlfs, R. J.; Springer, B. A.; Egeberg, K. D.; Sligar, S. G.; Tame, J.; Renaud, J.-P.; Nagai, K. *Nature* **1988,** *336,* 265-266.

(12) The eight helices of the globin fold are denoted A-H, starting with the N-terminus.

(13) Springer, B. A.; Egeberg, K. D.; Sligar, S. G.; Rohlfs, R. J.; Mathews, A. J.; Olson, J. S. *J. Biol. Chem.* **1989,** *264,* 3057-3060.

(14) Rohlfs, R. J.; Mathews, A. J.; Carver, T. E.; Olson, J. S.; Springer, B. A.; Egeberg, K. D.; Sligar, S. G. *J. Biol. Chem.* **1990,** *265,* 3168-3176. Carver, T. E.; Olson, J. S.; Smerdon, S. J.; Krzywda, S.; Wilkinson, A. J.; Gibson, Q. H.; Blackmore, R. S.; Dezz Ropp, J.; Sligar, S. G. *Biochemistry* **1991,** *30,* 4697-4705. Smerdon, S. J.; Dodson, G. G.; Wilkinson, A. J.; Gibson, Q. H.; Blackmore, R. S.; Carver, T. E.; Olson, J. S. *Biochemistry* **1991,** *30,* 6252-6260. Egeberg, K. D.; Springer, B. A.; Sligar, S. G.; Carver, T. E.; Rohlfs, R. J.; Olson, J. S. *J. Biol. Chem.* **1990,** *265,* 11788-11795.

(15) Carver, T. E.; Brantley, R. E., Jr.; Singleton, E. W.; Arduini, R. M.; Quillin, M. L.; Phillips, G. N., Jr.; Olson, J. S. *J. Biol. Chem.* **1992,** *267,* 14443-14450.

(16) Vyas, K.; Rajarathnam, K.; Lu, L. P.; Emerson, S. D.; La Mar, G. N.; Krishnamoorthi, R.; Mizukami, H. *J. Biol. Chem.* **1993,** *268,* 14826-14835.

(17) Pin, S.; Alpert, B.; Cortès, R.; Ascone, I.; Chiu, M. L.; Sligar, S. G. *Biochemistry* **1994,** *33,* 11618-11623.

(18) Lloyd, E.; Hildebrand, D. P.; Tu, K. M.; Mauk, A. G. *J. Am. Chem. Soc.* **1995,** *117,* 6434-6438.

(19) Adachi, S.; Magano, S.; Ishimori, K.; Watanabe, Y.; Morishima, I.; Egawa, T.; Kitagawa, T.; Makino, R. *Biochemistry* **1993,** *32,* 241-252.

(20) Dawson, J. H. *Science* **1988,** *240,* 433-439.

(21) Egeberg, K. D.; Springer, B. A.; Martinis, S. A.; Sligar, S. G.; Morikis, D.; Champion, P. M. *Biochemistry* **1990,** *29,* 9783-9791. Morikis, D.; Champion, P. M.; Springer, B. A.; Sligar, S. G. *Biochemistry* **1990,** *29,* 4791-4800.

(22) Hildebrand, D. P.; Burk, D. L.; Maurus, R.; Ferrer, J. C.; Brayer, G. D.; Mauk, A. G. *Biochemistry* **1995,** *34,* 1997-2005.

(23) DePillis, G. D.; Decatur, S. M.; Barrick, D.; Boxer, S. G. *J. Am. Chem. Soc.* **1994,** *116,* 6981-6982.

(24) Erman, J. E.; Vitello, L. B.; Miller, M. A.; Shaw, A.; Brown, K. A.; Kraut, J. *Biochemistry* **1993,** *32,* 9798-9806.

(25) Vitello, L. B.; Erman, J. E.; Miller, M. A.; Wang, J.; Kraut, J. *Biochemistry* **1993,** *32,* 9807-9818. Poulos, T. L. *Adv. Inorg. Biochem.* **1988,** *7,* 1-36.

(26) Goodin, D. B.; Mauk, A. G.; Smith, M. *Proc. Natl. Acad. Sci. U.S.A.* **1986,** *83,* 1295-1299. Sivaraja, M.; Goodin, D. B.; Smith, M.; Hoffman, B. M. *Science* **1989,** *245,* 738-740.

(27) Fitzgerald, M. M.; Churchill, M. J.; McRee, D. E.; Goodin, D. B. *Biochemistry* **1994,** *33,* 3807-3818.

(28) Spiro, T. G. In *Iron Porphyrins*; Lever, A. B. P., Gray, H. B., Eds.; Addison-Wesley: Reading, MA; Vol. 2, p 89.

(29) Vitello, L. B.; Erman, J. E.; Miller, M. A.; Mauro, J. M.; Kraut, J. *Biochemistry* **1992,** *31,* 11524-11535. Satterlee, J. D.; Erman, J. E.; Mauro, J. M.; Kraut, J. *Biochemistry* **1990,** *29,* 8797-8804. Goodin, D. B.; McRee, D. E. *Biochemistry* **1993,** *32,* 3313-3324. Ferrer, J. C.; Turano, P.; Banci, L.; Bertini, I.; Morris, E. K.; Smith, K. M.; Smith, M.; Mauk, A. G. *Biochemistry* **1994,** *33,* 7819-7829.

(30) McRee, D. E.; Jensen, G. M.; Fitzgerald, M. M.; Siegel, H. A.; Goodin, D. B. *Proc. Natl. Acad. Sci. U.S.A.* **1994,** *91,* 12847-12851.

(31) Sundaramoorthy, M.; Choudhury, K.; Edwards, S. L.; Poulos, T. L. *J. Am. Chem. Soc.* **1991,** *113,* 7755-7757.

(32) Choudhury, K.; Sundaramoorthy, M.; Hickman, A.; Yonetani, T.; Woehl, E.; Dunn. M.; Poulos, T. L. *J. Biol. Chem.* **1994,** *269,* 20239-20249.

(33) Raphael, A. L.; Gray, H. B. *J. Am. Chem. Soc.* **1991,** *113,* 1038-1040.

(34) Wallace, C. J. A.; Clark-Lewis, I. *J. Biol. Chem.* **1992,** *267,* 3852-3861.

(35) Raphael, A. L.; Gray, H. B. *Proteins* **1989,** *111,* 766-767.

(36) Bren, K. L.; Gray, H. B. *J. Am. Chem. Soc.* **1993,** *115,* 10382-10383.

(37) Lu, Y.; Casimiro, D. R.; Bren, K. L.; Richards, J. H.; Gray, H. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993,** *90,* 11456-11459.

(38) Takano, T.; Dickerson, R. E. *J. Mol. Biol.* **1981,** *153,* 95-115.

Chapter 2

Semisynthesis, Characterization, and Ligand-Binding Properties of Horse Heart Ala80cytochrome *c* and Phe67Ala80cytochrome *c*

Background

Cytochrome *c*

Cytochromes *c* are heme-containing redox proteins involved in respiration or photosynthesis.[1] Cytochromes *c* are distinguished from other cytochromes by the covalent means of heme attachment to the polypeptide; the vinyl groups of the protoporphyrin IX group in cytochromes *c* are replaced by thioether bridges to two (or, in a few cases, one) cysteine residues (Figure 2.1). In all cytochromes *c*, the iron in the heme (or hemes) cycles between the oxidized Fe(III) state and the reduced Fe(II) state, resulting in a net charge at the iron site of +1 and 0, respectively, due to coordination by the dianionic heme. The UV/vis absorption spectra reflect these changes, with the intense Soret (or γ) band increasing in intensity and red-shifting upon reduction and the lower-energy α, β bands splitting (Figure 2.2).[2] These highly-allowed $\pi \rightarrow \pi^*$ absorption bands give cytochromes their characteristic deep red color.

The most familiar of the cytochromes *c* are those found in the mitochondria of eurkaryotes. These proteins, the class I cytochromes *c*, serve the vital biological function of shuttling electrons from the cytochrome bc_1 complex to cytochrome *c* oxidase in the mitochrondrial respiratory electron transport chain. Because cyt *c* plays a crucial role in this vital process that results in ATP production, it is ubiquitous in nature; over 90 varieties have been sequenced and many more have been isolated.[1a] The wide availability of cyt *c*, coupled with its stability, water-solubility, and intense absorption features, has made it an irresistible subject of study by scientists in many fields. The resulting depth of knowledge on its structure and function facilitates studies in areas such as protein folding, protein engineering, electron transfer, and spectroscopy.

Mitochondrial cytochromes *c* consist of one heme *c* and a single polypeptide chain with 103–113 amino acids. The reduction potentials of mitochondrial cytochromes *c* are quite high, lying between +200 and +390 mV vs. NHE. The three-dimensional structure of a number of cytochromes *c* is known and, despite wide sequence variability,

Figure 2.1 X-ray crystal structure (PDB identifier 1HRC) of the cyt *c* heme, the thioether linkages to Cys14 and Cys17, and the axial ligands His18 and Met80.

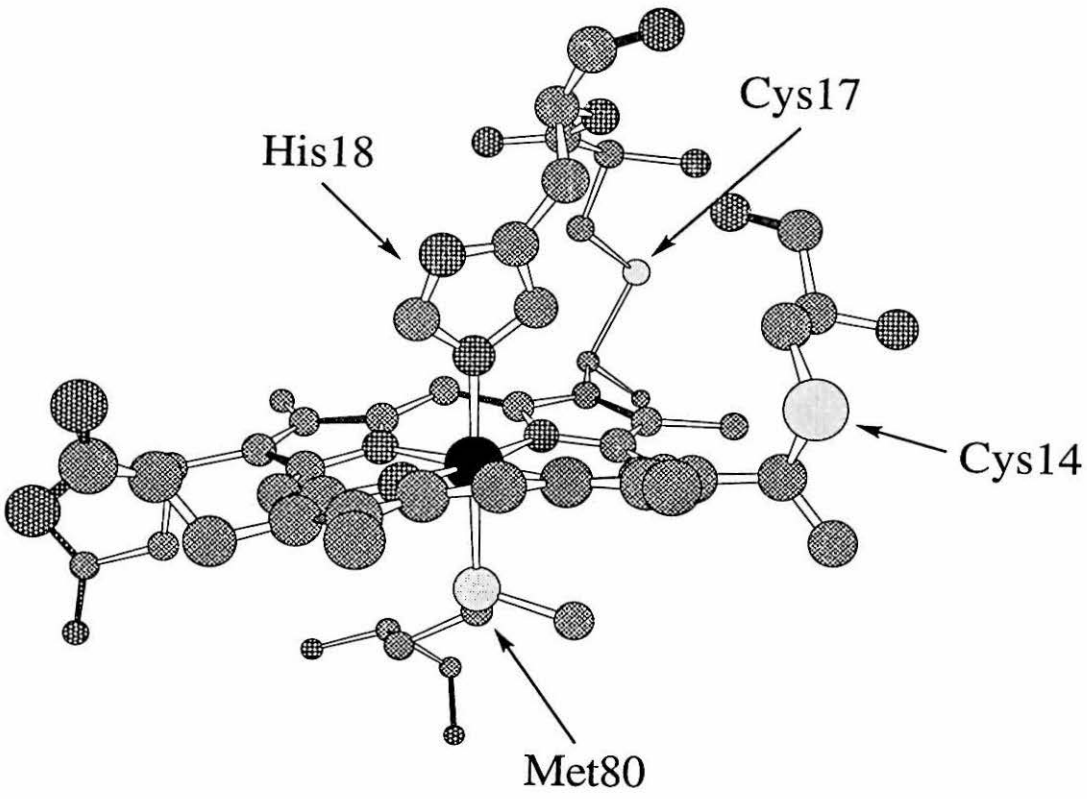

His18
Cys17
Cys14
Met80

Figure 2.2 UV/vis absorption spectra of oxidized (dotted line) and reduced (solid line) native horse heart cyt *c* in 25 mM NaP_i (pH 7.0).

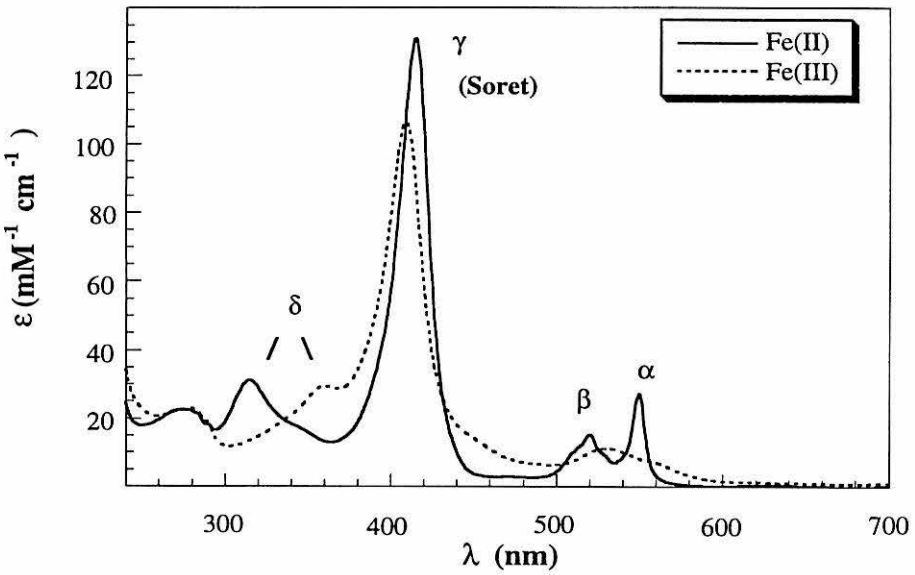
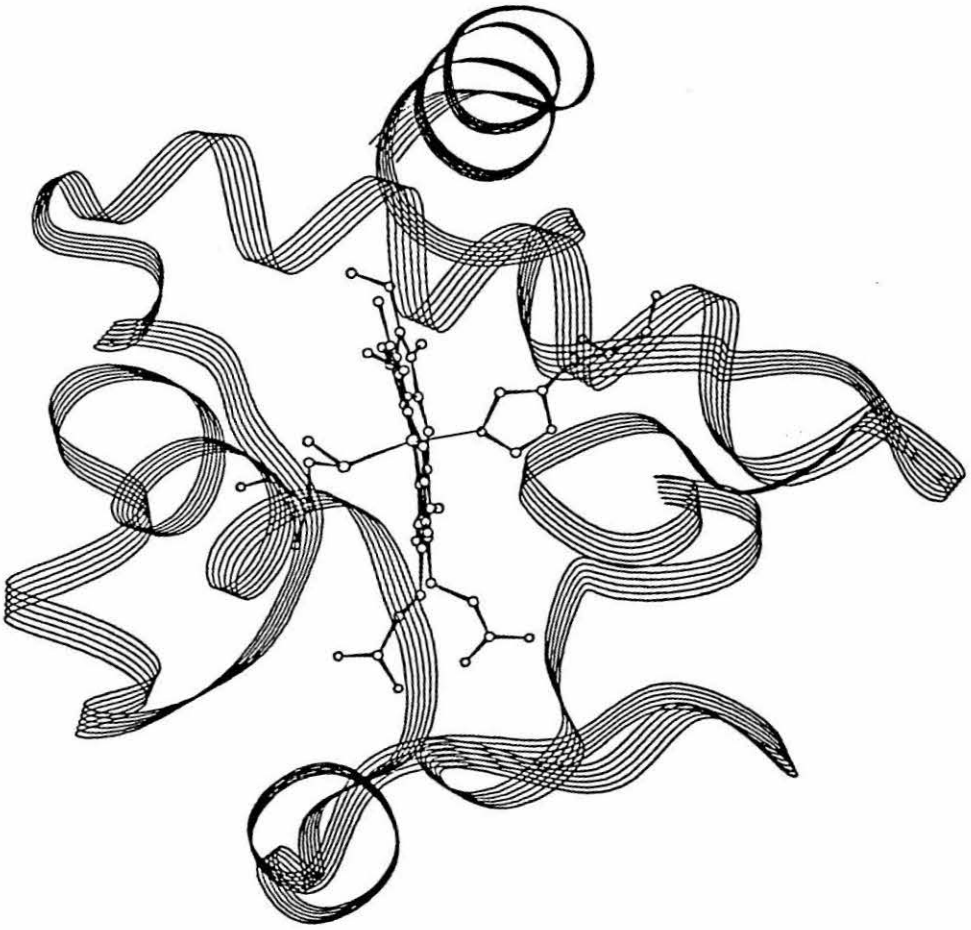

γ
(Soret)
Fe(II)
Fe(III)
δ
β
α
120
100
80
60
40
20
ε (mM-1 cm-1)
300
400
500
600
700
λ (nm)

Figure 2.3 X-ray crystal structure of ferric horse heart cyt *c* (PDB identifier 1HRC). The ribbon illustrates the cytochrome fold. The heme group and the side chains of the two heme axial ligands, His18 and Met80, are shown.



they display remarkably similar folds.[1,3] The typical class I cytochrome fold (Figure 2.3) encloses most of the heme, leaving part of one heme edge exposed. The heme iron is axially coordinated by two highly conserved residues, His18 and Met80. The axial His binds the heme with its ring plane aligned with the heme α and γ meso carbons; it has been shown that the heme-attaching loop (residues 14–18) constrains the His in this orientation.[4] Class I cyts *c* are ~ 45% α-helical and begin and end in α-helices that interact with each other. Three other α-helices are also found in cyt *c*. Coiled structure, loops, and 6 type II 3_{10} bends make up the rest of the fold, giving cyt *c* a compact structure. The heme pocket is very hydrophobic, typically with many leucine residues surrounding the heme. A number of conserved aromatic residues are also found near the heme. The hydrophobic interactions in the core of cyt *c*, coupled with an extensive hydrogen-bonding network, create an extremely stable structure; most class I cytochromes *c* do not unfold until heated above 60 °C.[1a] On the surface, cytochrome *c* has many lysine residues, resulting in a significant overall positive charge. Many of these lysines are found near the exposed heme edge and this positive patch is thought to be important in its interactions with its redox partners.[1]

Semisynthesis

Semisynthesis is the reconstitution of a protein fragment with a synthetic peptide to form the intact protein. This technique has been utilized to probe protein structure and function in a number of systems.[5] Horse heart (h. h.) cyt *c* is particularly amenable to the semisynthetic methodology. In the early 1970's, it was discovered that cyanogen bromide (CNBr) cleavage fragments of cyt *c* (cyt *c*(1–65) and (66–104)) are able to form a covalently-bound complex with properties similar to those of the native protein.[6] This has been used to probe structure-function relationships in cyt *c* by modifying the 66–104 fragment prior to protein reconstitution and by producing chimeric proteins.[7] However, it was not until the development of solid-phase peptide synthesis technology[8] that the

semisynthetic method, in which a native protein fragment is religated with a synthetic peptide,[9] became feasible.

The semisynthetic methodology begins with the preparation of the cyt *c*(1–65) fragment by incubation of h. h. cyt *c* with CNBr, which specifically cleaves after methionine residues. Because h. h. cyt *c* has two methionine residues (at positions 80 and 65, Figure 2.4), CNBr cleavage yields five fragments which can be separated by size-exclusion or ion-exchange chromatography.[6a] The 1–65 and 1–80 fragments contain the heme, which is linked *via* residues 14, 17, and 18. The CNBr cleavage reaction (Figure 2.5) leaves a homoserine (Hse) lactone in place of the C-terminal methionine, thus providing a C-terminus activated for coupling with the N-terminus of a peptide.

The mechanism of the religation reaction is proposed to begin with complex formation between cyt *c*(1–65) and the 66–104 peptide (Figure 2.6). If the 1–65 heme-containing fragment and 39-mer peptide are combined under aerobic conditions, a complex with properties different from the native protein is formed (*i.e.*, the Met → Fe(III) charge transfer band at 695 nm is absent).[6] Reduction of this complex triggers a conformational change that brings the Hse lactone and peptide N-terminus into close proximity so that the peptide bond can be formed, giving Hse65cyt *c*. Replacement of Met65 with Hse is not expected to alter the structure or stability of cyt *c* as it is a surface residue with significant variability among species.[1a] The insignificance of the Met65Hse mutation has been demonstrated experimentally through extensive spectroscopic and functional studies of semisynthetic cyt *c*. The absorption spectrum (including the 695 nm band in the oxidized state), CD spectrum, and reduction potential of Hse65cyt *c* are indistinguishable from those of native cyt *c*.[6,7a,9a] The NMR spectra of both oxidized and reduced semisynthetic cyt *c* are superimposable with native.[10] In addition, the native and semisynthetic proteins have equal reactivity to anti-cyt *c* antibodies,[9a] and similar biological activity in cyt *c*-depleted mitochondria.[7a]

Figure 2.4 The amino acid sequence of h. h. cyt *c*. The heme axial ligands (positions 18 and 80) are in bold type and the Cys thioether linkages (positions 14 and 17) are underlined.

1	Ac-Gly	Asp	Val	Glu	Lys	Gly	Lys	Lys	Ile	Phe
11	Val	Gln	Lys	Cys	Ala	Gln	Cys	**His**	Thr	Val
21	Glu	Lys	Gly	Gly	Lys	His	Lys	Thr	Gly	Pro
31	Asn	Leu	His	Gly	Leu	Phe	Gly	Arg	Lys	Thr
41	Gly	Gln	Ala	Pro	Gly	Phe	Thr	Tyr	Thr	Asp
51	Ala	Asn	Lys	Asn	Lys	Gly	Ile	Thr	Trp	Lys
61	Glu	Glu	Thr	Leu	Met	Glu	Tyr	Leu	Glu	Asn
71	Pro	Lys	Lys	Thr	Ile	Pro	Gly	Thr	Lys	**Met**
81	Ile	Phe	Ala	Gly	Ile	Lys	Lys	Lys	Thr	Glu
91	Arg	Glu	Asp	Leu	Ile	Ala	Tyr	Leu	Lys	Lys
101	Ala	Thr	Asn	Glu-OH						

Figure 2.5 Mechanism of CNBr cleavage after Met residues. The homoserine lactone and carboxylate products are indicated.

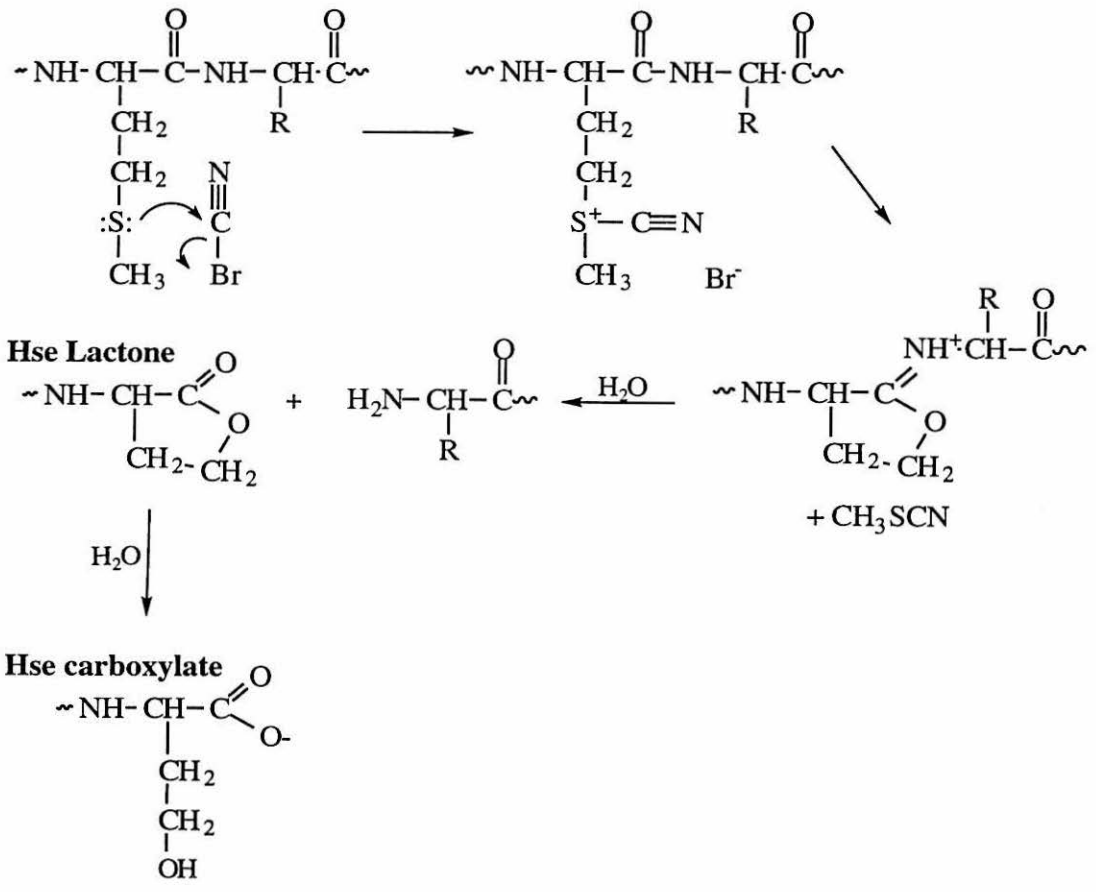
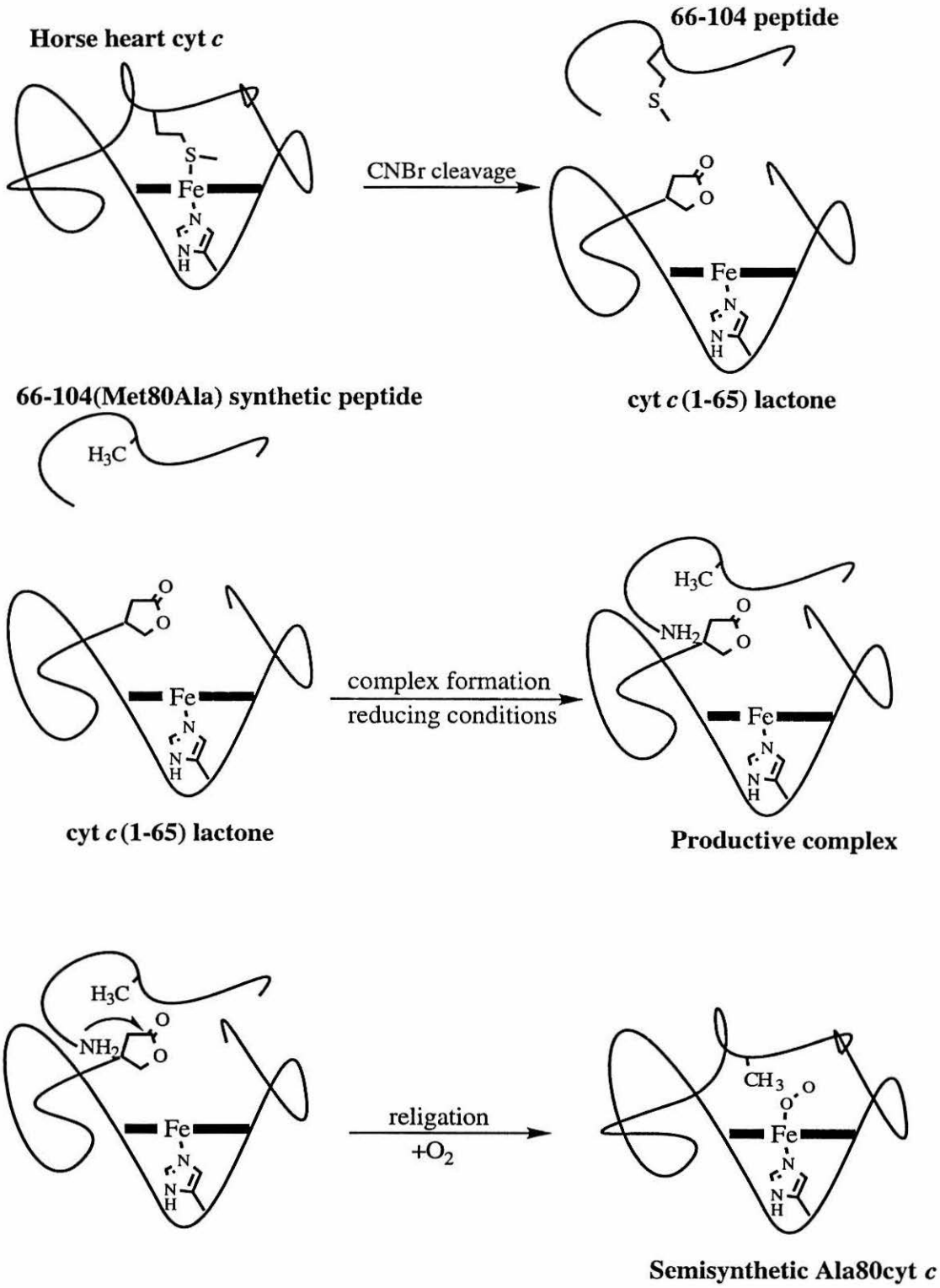

Hse Lactone
Hse carboxylate
H2O
+ CH3SCN

Figure 2.6 Illustration of the semisynthesis of Ala80cyt *c*. First, native h. h. cyt *c* is cleaved with CNBr to produce cyt *c*(1–65) lactone which is incubated with a synthetic 66-104 peptide with Ala substituted for Met80. Cyt *c*(1–65) lactone and the peptide form a productive complex, bringing the peptide N-terminus into proximity with the cyt *c* lactone moiety, allowing religation of the peptide bond.

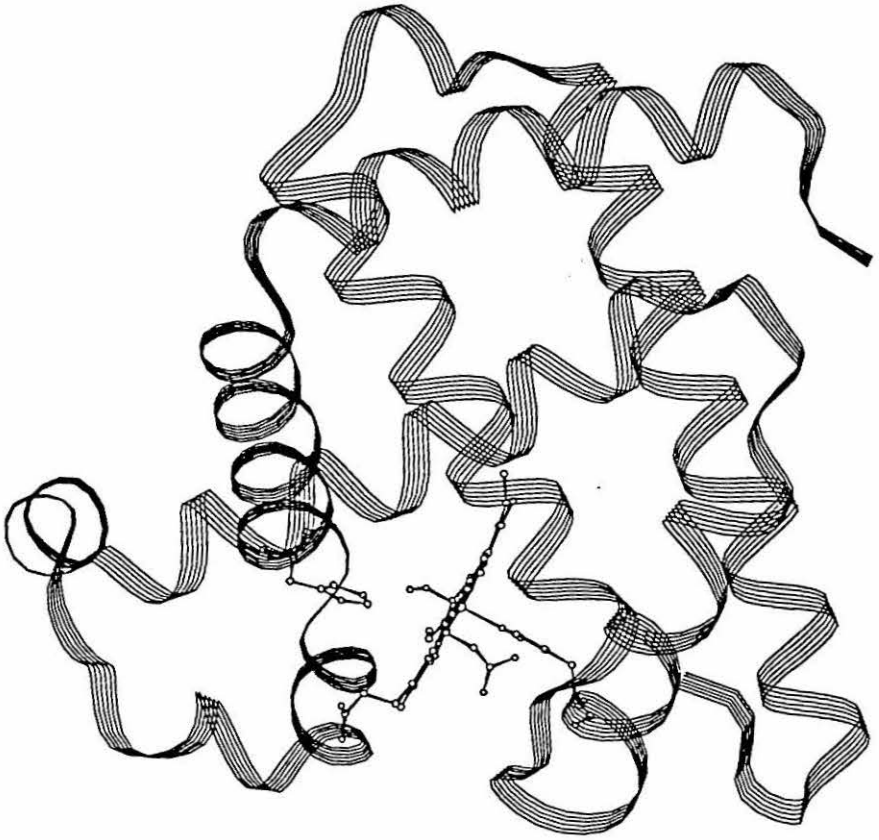

Horse heart cyt c
66-104 peptide
CNBr cleavage
66-104(Met80Ala) synthetic peptide
cyt c (1-65) lactone
complex formation
reducing conditions
cyt c (1-65) lactone
Productive complex
religation
+O2
Semisynthetic Ala80cyt c

Semisynthesis was chosen as a technique for the preparation of cyt *c* variants with dramatically different properties from native cyt *c* (*i.e.*, altered heme ligation) because, prior to a recent breakthrough,[11] it was not possible to express non-functional cyts *c* by recombinant means. Semisynthesis remains a powerful technique, in particular for the incorporation of non-encoded amino acids into the 66–104 region of cyt *c*. Using semisynthesis, *p*-fluorophenylalanine has been substituted for Tyr67,[7b,9h] unnatural amino acids have been introduced at positions 66,[7c] 78,[9b] and 82,[9j] and Met80 has been replaced with a variety of unnatural amino acids.[7d,9g] A variant with enhanced electron-transfer properties has been produced by introduction of a 2,2'-bipyridylalanine amino acid into position 72.[9i] The limitations of the semisynthetic technique include the fact that no changes that significantly alter protein structure or that affect complex formation can be introduced and that it is only possible to introduce mutations into the 66–104 portion of the protein. Nonetheless, semisynthesis remains the only method presently available for preparing multi-milligram quantities of cyt *c* containing unnatural amino acids. Semisynthesis also provides an alternative route to some mutants which cannot be expressed by recombinant methods.

Ligand Binding to Myoglobin

Myoglobin (Mb) is a respiratory heme protein found in the smooth or striated muscles of animals.[12] Like cyt *c*, it is small (18 kDa), monomeric, and highly soluble in water. The Mb polypeptide is highly α-helical, with 77% of its amino acids (in sperm whale Mb) located in the 8 α-helical regions denoted A–H (Figure 2.7). The heme in Mb is accommodated by a hydrophobic pocket near the surface of the protein, with part of one heme edge exposed to solvent.[13] The heme iron is bound by the proximal His (His F8 in helical notation) toward which it moves slightly out of the porphyrin plane. The heme propionates have been shown to be important for positioning the heme within the heme cleft, forming hydrogen bonds with amino acids His FG3, Ser F6, and Arg CD3.[14]

Figure 2.7 X-ray crystal structure of O_2-Fe(II)Mb from sperm whale (PDB identifier 1MBO). The ribbon illustrates the highly helical secondary structure. The heme and the proximal and distal His side chains are also shown.



Numerous van der Waals contacts between the heme and the polypeptide assist in securing the heme in its place. A number of aromatic rings near the heme lie parallel to the heme plane and are thought to interact with the heme through π-stacking interactions.[12,13]

Mb performs the difficult task of reversibly binding dioxygen. Only the ferrous state of heme has any affinity for dioxygen. However, ferrous heme in solution rapidly autoxidizes to the ferric state (denoted "met") in the presence of dioxygen.[15] The globin therefore is essential for allowing the heme to form a stable dioxygen complex.[16] The stability of the dioxygen-heme complex can be considered in two ways: kinetic and thermodynamic. The kinetic stability is exhibited by the autoxidation rate, while thermodynamic stability is seen in the dioxygen binding constant. The resistance of myoglobin to autoxidation can be attributed in part to the tight binding of the heme to the polypeptide to prevent the dimerization and oxidation of hemes to form μ-oxo-dimers,[17] forcing Mb to find an alternative autoxidation pathway. The autoxidation mechanism proposed for Mb (Figure 2.8) requires the release of superoxide or dioxygen from the heme, followed by the binding of water;[18] the presence of the polypeptide barrier between the heme and the solvent slows these processes. The thermodynamic stability of oxy-Mb, as seen in low (micromolar) dioxygen dissociation constants, is afforded by the stabilization of the polar Fe(II)-O_2 unit and its resonance form, Fe(III)-O_2^-, by hydrogen-bonding between the distal His E7 and the heme-bound dioxygen (Figure 2.9).[19,20] The low O_2 affinity exhibited by distal His mutants demonstrates the functional importance of this residue (Table 2.1).[20]

The Mb heme binds a number of ligands other than dioxygen.[12] As mentioned above, ferric Mb binds water. This "metMb" form is high-spin, with the heme iron 0.3 Å out of the plane of the heme. The water deprotonates at pH ~ 9 to give a low-spin hydroxide-bound heme. Azide and cyanide, among other ligands, form stable complexes with Fe(III)Mb. In the ferrous state and in the absence of dioxygen, Mb is 5-coordinate

Figure 2.8 Schematic representation of the mechanism of Mb autoxidation. At high $[O_2]$, the dominant pathway is through direct dissociation of superoxide (HO_2) while at low $[O_2]$, autoxidation occurs by long-range electron transfer from Fe(II)Mb to a dissociated dioxygen molecule. His-H and H-His represent distal His tautomers. Adapted from reference 18.

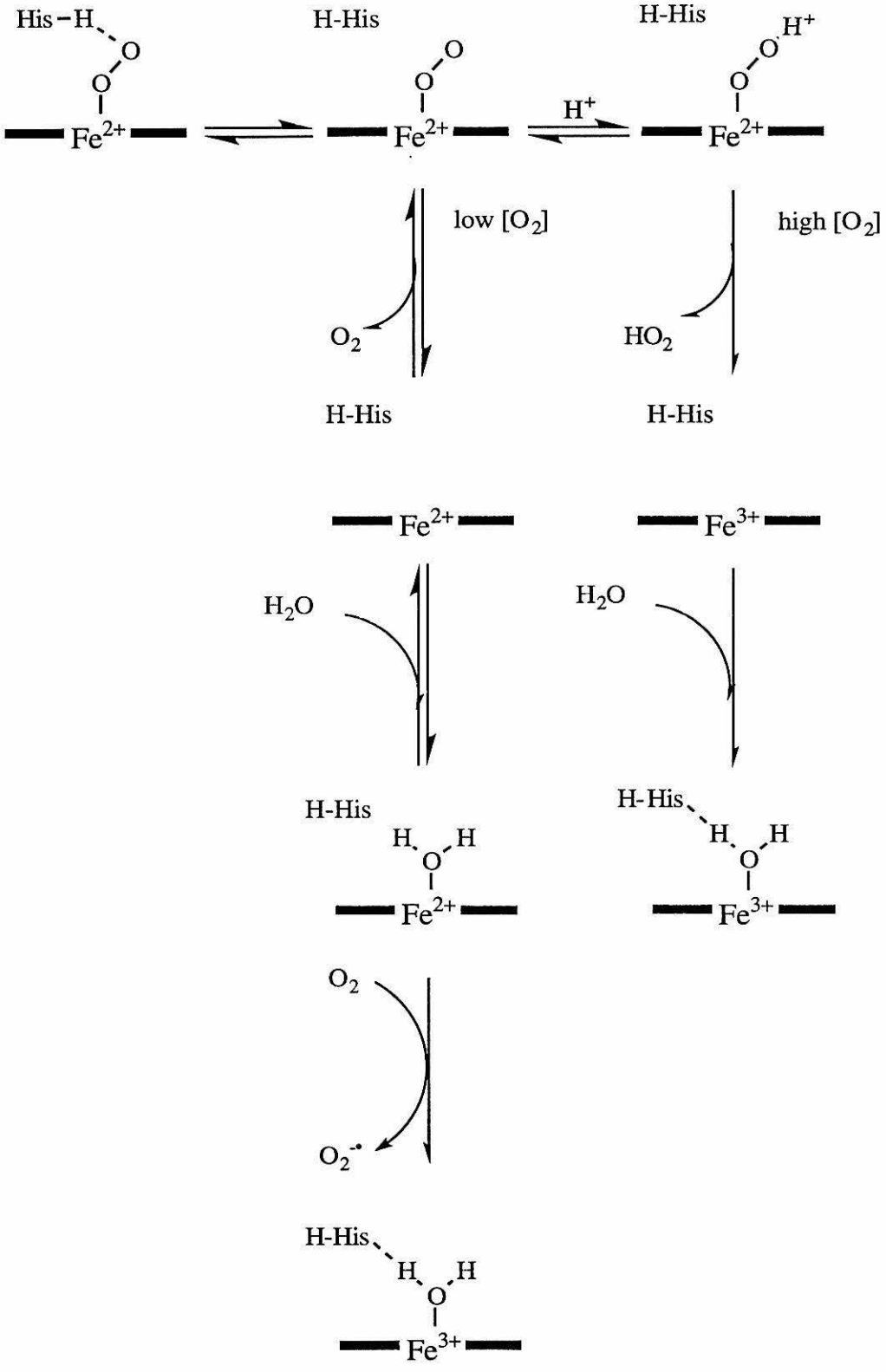
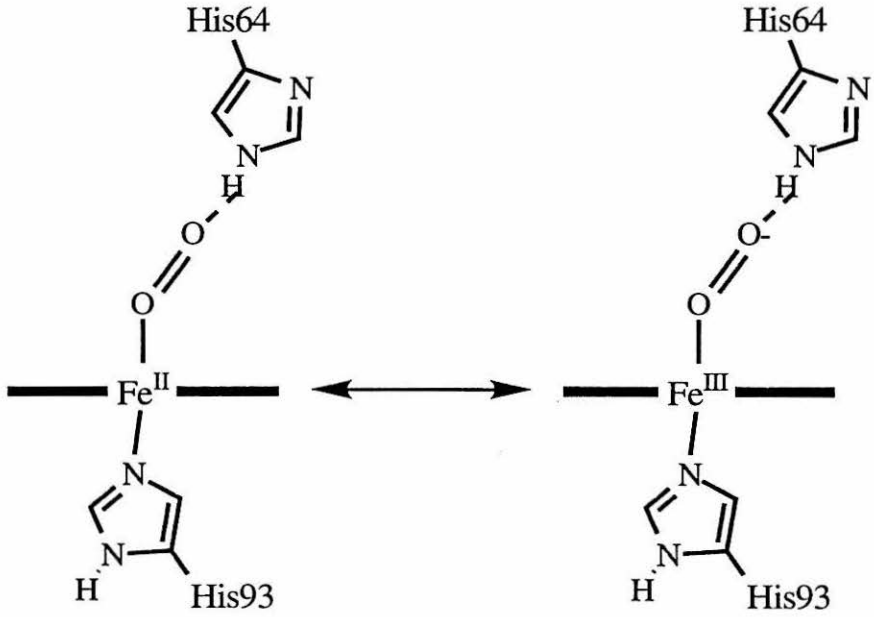

His–H
H-His
H-His
H+
Fe2+
Fe2+
Fe2+
low [O2]
high [O2]
O2
HO2
H-His
H-His
Fe2+
Fe3+
H2O
H2O
H-His
H-His
Fe2+
Fe3+
O2
O2-
H-His
Fe3+

Figure 2.9 Representation of a hydrogen-bonding interaction between a His side chain and a heme-dioxygen unit.


His64
N
N
H
O
O
Fe^II
N
N
H
His93
His64
N
N
H
O-
O
Fe^III
N
N
H
His93

and high-spin. CO, NO, and isocyanides form very stable complexes with Fe(II)Mb, all resulting in low-spin species. Free heme in solution, in fact, shows a ~ 1000x preference for binding CO over O_2.[15] With the amount of CO present in the body as a result of heme metabolism, ~ 20% of the Mb in the body would be occupied by CO rather than O_2 if Mb showed a similar preference for CO.[15,21] However, the polypeptide is able to destabilize CO binding so that K_{CO}/K_{O_2} is only ~ 30 for Mb,[12,15,16] a number low enough to allow Mb to function effectively *in vivo*. One hotly-debated proposal for the mechanism of CO-Fe(II)Mb destabilization is that the distal His forces the CO to bend and/or tilt from its favored position normal to the heme plane, reducing the binding constant.[16,22]

Table 2.1

Effect of distal histidine mutations on ligand-binding properties of sperm whale Mb

Mutant	K_{O_2} x 10^{-6} M^{-1}	K_{CO} x 10^{-6} M^{-1}	K_{CO}/K_{O_2}	k_{autox}[a] h^{-1}	k_{autox} mutant/WT	refs
native	1.2	27	23	0.36	1	20a,b
Gly	0.088	150	1,700	39	110	20a,b,c
Val	0.011	150	13,600	36	100	20a,b
Leu	0.023	1100	48,000	**	**	20a,b
Met	0.045	200	4,400	16	50	20a,b
Phe	0.0074	83	11,000	16	45	20a,b
Gln	0.18	82	455	**	**	20a
Arg	0.090	400	4,400	18	53	20b
Asp	*	85	*	120	350	20b
Cys	**	**	**	43	120	20b
Lys	**	**	**	30	87	20b
Thr	**	**	**	42	120	20b
Tyr	*	6.6	*	>100	>300	20b

[a]at 37 °C.

*autoxidation is too fast to determine K_{O_2} for these mutants.

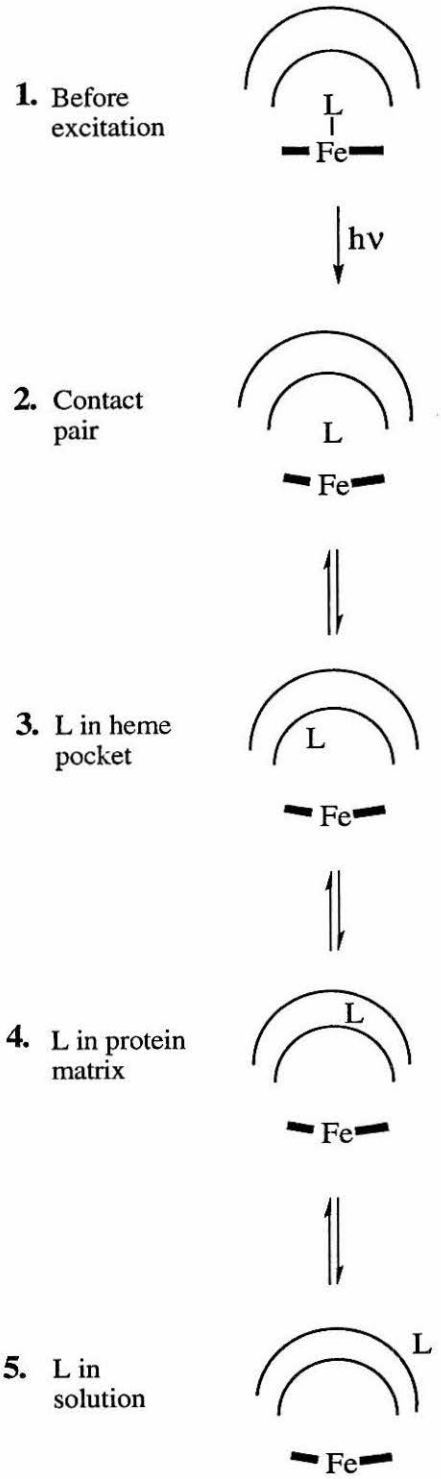
**not given.

The photolability of ligands from ferrous hemes has been utilized to probe the ligand-binding mechanism of heme proteins, Mb in particular.[23-27] The kinetics of photodissociation and subsequent recombination of small exogenous ligands (most commonly CO, O_2, and NO) from myoglobin and other heme proteins has been a subject of study for many years.[23] With the application of ultrafast (nanosecond to sub-picosecond) transient absorption and Raman spectroscopy, it was found that the ligand-binding process is more complex than first thought.[24-26] These experiments have led to a greater understanding of heme and ligand dynamics in proteins.

Hexacoordinate ferrous heme proteins go through a number of states after photoexcitation, from an excited state ligand-bound form to a fully-dissociated product with the ligand outside of the protein matrix. The following scheme is proposed[25] (refer to Figure 2.10): Absorption of a photon results in rupture of the iron-ligand bond. If the ligand rebinds from this contact pair intermediate, geminate (unimolecular) recombination occurs that is seen on a picosecond or sub-picosecond timescale. Alternatively, the ligand may diffuse through the heme pocket or into the protein matrix, from where it may rebind on a nanosecond timescale. The third possibility is that the ligand may diffuse out of the protein and into the solvent. Bimolecular recombination kinetics are observed from this state on a timescales typically in the millisecond range. Conformational changes of the heme group accompany ligand binding, with the heme moving from the high-spin bowed conformation with the iron out of the heme plane into the planar low-spin ligand-bound conformation as the ligand binds the heme iron.[26]

The expression of myoglobins with mutations in the heme pocket has allowed the study of the effects of distal pocket structure on ligand-binding kinetics and therefore on heme protein dynamics. Ligand-binding studies on variants with residues replacing the distal His have led to the conclusion that the polar nature of His64 is the primary governing force for association rate constants for CO and O_2 binding.[20d] The kinetics of

Figure 2.10 Schematic representation of the intermediates of ligand photodissociation and binding to Fe(II)Mb. The area within the inner semicircle represents the distal heme pocket, and the outer semicircle represents the protein/solution interface. After photoexcitation, the iron-ligand bond breaks within 300 femtoseconds to form the contact pair, from which rebinding may occur on the femtosecond or picosecond timescale. Alternatively, the ligand may move through the heme pocket, into the protein matrix, and/or into the solvent and may rebind from any of these intermediate states.


1. Before excitation
L
Fe
hν
2. Contact pair
L
Fe
3. L in heme pocket
L
Fe
4. L in protein matrix
L
Fe
5. L in solution
L
Fe

ligand binding to position 29 mutants indicate that Leu29 acts as a barrier which restricts ligand movement in the heme pocket.[27]

Materials and Methods

General

Water used for all buffer solutions was purified by passing distilled water through a Barnstead Nanopure purification system (No. 2794) equipped with two ion-exchange filters, one organic-removal filter and one 0.22 μm filter. Specific resistance of all water samples was > 18 MΩ-cm. Sodium phosphate (NaP_i) and N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES) were of analytical grade. Sodium dithionite (~ 85% purity) was obtained from Sigma and kept over Dri-rite. The highest grade available (purity > 95%) of h. h. cyt *c* was purchased from Sigma. Horse heart myoglobin (purity 95–100%) was purchased from Sigma. CNBr (> 97%) was obtained from Aldrich. K[Co(EDTA)] was prepared as described,[28] with sodium carbonate substituted for barium carbonate.

Protein solutions were stored at 4 °C (< 2 weeks) or at -80 °C after flash-freezing in liquid nitrogen (> 2 weeks). Protein solutions were repurified if kept at 4 °C for over 2 weeks. Concentration of protein solutions was performed at 4 °C (unless noted otherwise) by ultrafiltration (for solutions of > 10 mL) using cells (200 mL, 50 mL, or 10 mL) and YM-3 membrane filters (3000 MW cutoff) from Amicon. For smaller volumes, Centricon 10 microconcentrators (10,000 MW cutoff) from Amicon were used in a Sorvall RC-5 superspeed refrigerated centrifuge with a SS-34 rotor spinning at 5000 rpm. pH measurements were made at room temperature with a Beckman model Φ32 pH meter, using a Beckman combination electrode or an Ingold microelectrode. Calibration was done with VWR buffer standards.

Semisynthesis of Cytochrome *c* Variants

Cytochrome c Purification

Horse heart cyt *c* was always purified by FPLC before use. The Pharmacia FPLC apparatus consisted of two P-500 pumps, a MV7 valve, a UV-M monitor with a mercury lamp with 280 and 405 nm filters, a FRAC-100 fraction collector, an LCC-500 controller, and a REC-482 chart recorder. Samples were loaded *via* a 10 mL superloop. Mono S HR (high-resolution cation exchange) 16/10 or 10/10 columns were used. Buffers and samples were filtered through 2.2 μm filters immediately prior to application to the column. All columns were equilibrated and cleaned according to the manufacturer's instructions.

Horse heart cyt *c* (200–400 mgs) was dissolved in 20–50 mL buffer (25 mM NaP_i, pH 7.0) along with ~ 10x excess K[Co(EDTA)] and gently stirred at 4 °C for at least 12 hours prior to purification to assure that all of the protein was in the oxidized form. Using the Mono S HR 16/10 column, 20–40 mgs of protein could be purified per loading. Cyt *c* in 25 mM NaP_i (pH 7.0) was separated from impurities (mostly believed to be deamidated forms of cyt *c*) by application of a gradient from 0 to 16 mM NaCl in 75 mL followed by a gradient from 16 to 20 mM NaCl in 120 mL. $Co(EDTA)^-$ eluted in the void volume and Fe(III)cyt *c* eluted at 18 mM NaCl. Protein concentrations were determined using molar extinction coefficients reported in the literature for oxidized h. h. cyt *c*: $\varepsilon_{410} = 106.1\ mM^{-1}cm^{-1}$, $\varepsilon_{528} = 11.2\ mM^{-1}cm^{-1}$.[29]

Preparation of Cytochrome c(1–65)

Cyt *c*(1–65) was prepared as reported in the literature with slight modifications.[9c] A solution of purified cyt *c* was successively concentrated by ultrafiltration and diluted with H_2O until the calculated ionic strength was $< 10^{-5}$ M and the sample volume was < 5 mL. The concentrated protein solution was diluted with 88% formic acid to give a 15 mg/mL protein solution in 70% formic acid. The color was observed to change from

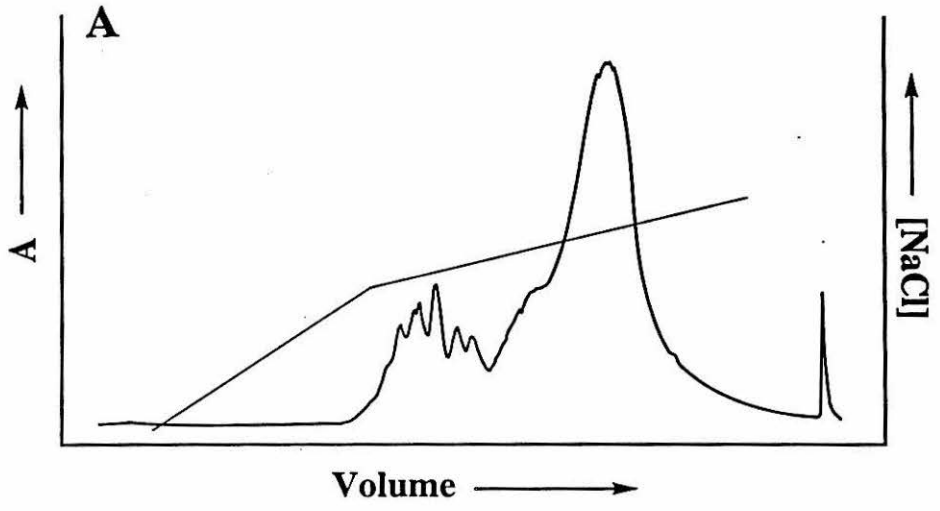
deep red to brown; under these conditions, the protein is denatured and the heme high-spin. A 50x excess of CNBr was added to the mixture in a round-bottom flask and the flask sealed with a greased ground-glass stopper. The reaction was stirred gently in the dark for ~ 40 hours and then quenched by dilution with ~ 10x excess H_2O followed by ultrafiltration. Ultrafiltration was performed at room temperature in the hood. The filtrate was collected in a solution with 100x excess $AgNO_3$ with respect to CNBr to precipitate solid AgCN which is not volatile as is CNBr. The solution was concentrated and diluted with 25 mM NaP_i (pH 7.0) buffer until the filtrate was at pH 7 (5–6 cycles).

Purification of cyt *c*(1–65) was performed by application of the neutralized reaction mixture to a Mono S HR 16/10 column. The protein fragments in 25 mM NaP_i (pH 7.0) were separated by application of a gradient from 0–25 mM NaCl over 100 mL followed by a gradient from 25–28 mM NaCl over 100 mL (Figure 2.11A). The largest peak (monitoring at 280 nm) was cyt *c*(1–65), which eluted at 30–32 mM NaCl. Because the Hse lactone hydrolyzes with time, a second purification was carried out to separate cyt *c*(1–65) lactone from cyt *c*(1–65) carboxylate (Figure 2.11B). The above purification protocol was used, except that the second part of the gradient was run over 150 mL. Cyt *c*(1–65) carboxylate appears as a shoulder on the major peak, eluting at lower ionic strength than cyt *c*(1–65) lactone as expected. Samples were used within three days of purification.

Peptide Synthesis

Peptides were synthesized and purified at the Caltech Biopolymer Synthesis and Analysis Resource Center under the direction of Susannah Horvath. Two 39-mer peptides corresponding to residues 66–104 of h. h. cyt *c* were constructed: one with Ala at position 80, and a second with Ala at position 80 and Phe at position 67. An ABI 430A synthesizer and t-BOC protocols were utilized in the synthesis.[8] Peptides were desalted and purified by reversed-phase HPLC on a C-8 column (Figure 2.12) and stored

Figure 2.11 Typical cation-exchange (FPLC Mono S HR 16/10) chromatograms for the (A) first and (B) second purifications of cyt *c*(1–65) in 25 mM NaP_i (pH 7.0). The thick lines are the solution absorbance measured at 280 nm on the 2.0 absorbance scale and the thin lines are the concentration of NaCl. In both (A) and (B), the major band is the cyt *c*(1–65) lactone product.


A
A
[NaCl]
Volume

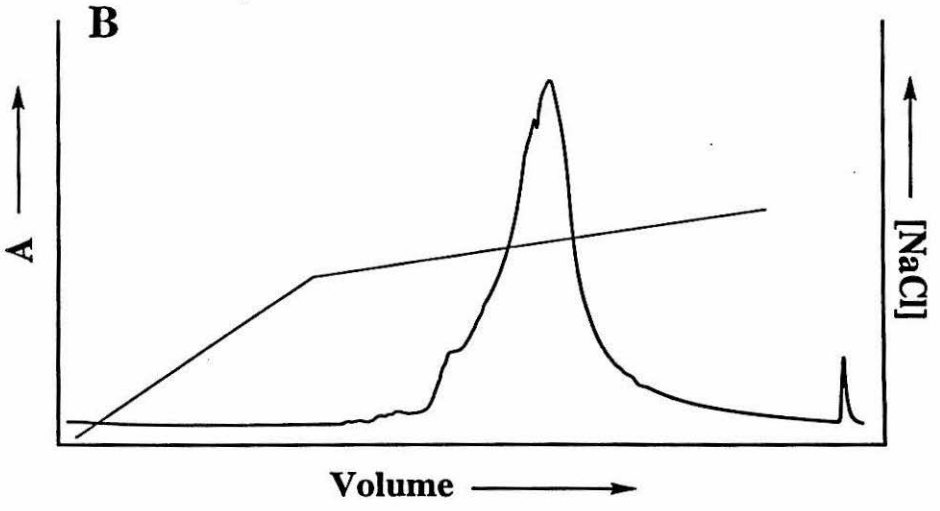
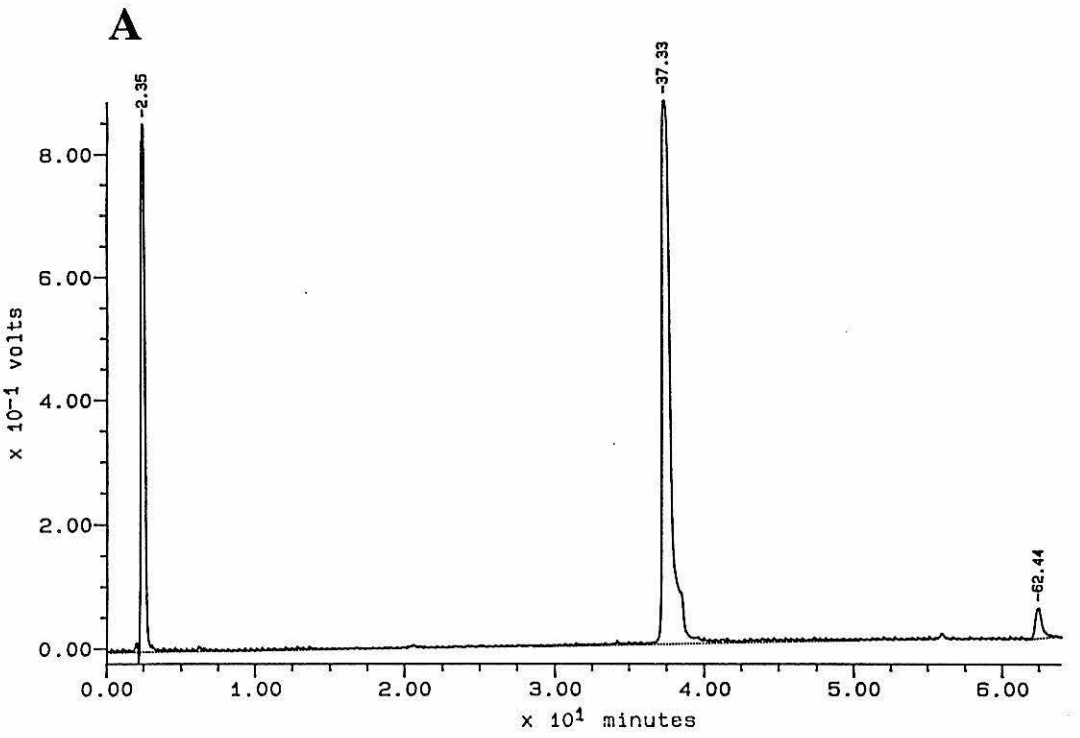
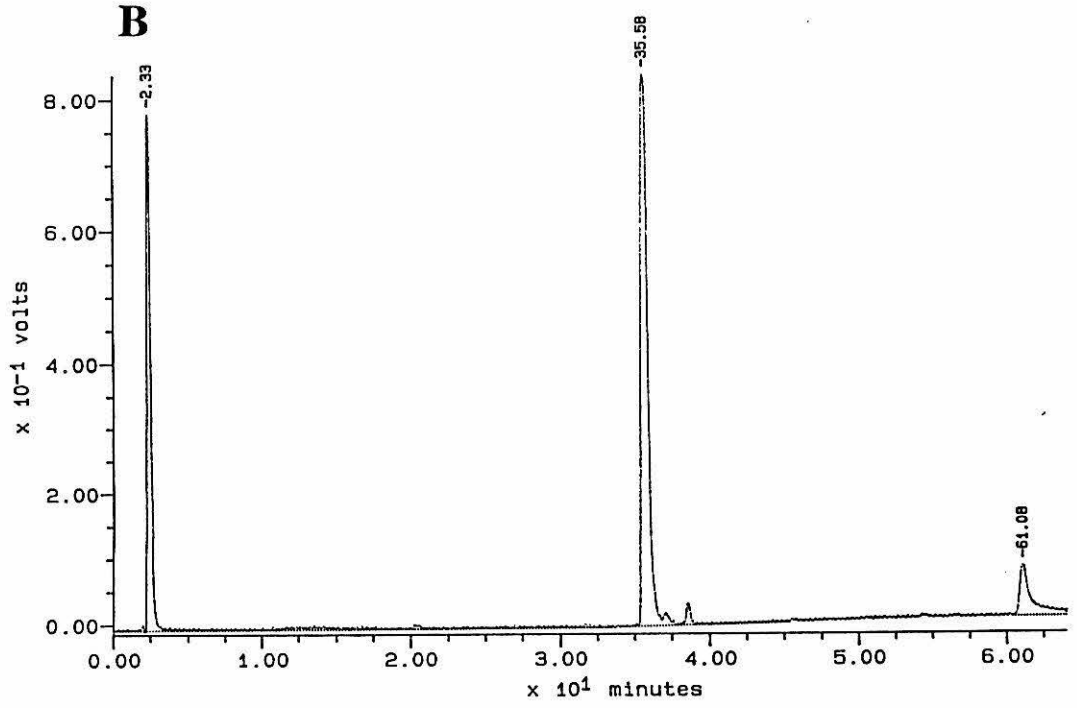

B
A
[NaCl]
Volume

Figure 2.12 Reversed-phase (C−8 column) HPLC trace of the final purification of the (A) 66−104(Ala80) and (B) 66−104(Phe67Ala80) peptides. The large peak eluting early is due to small-molecule impurities and reagents and the large peaks eluting at (A) 37 and (B) 35 minutes correspond to the pure peptide products.


A
2.35
37.33
62.44
8.00
6.00
4.00
2.00
0.00
x 10-1 volts
0.00
1.00
2.00
3.00
4.00
5.00
6.00
x 10¹ minutes


B
2.33
35.58
61.08
8.00
6.00
4.00
2.00
0.00
x 10-1 volts
0.00
1.00
2.00
3.00
4.00
5.00
6.00
x 10¹ minutes

lyophilized over desiccant. The peptides were characterized by elemental analysis, amino acid composition, sequencing, and mass spectroscopy.

Protein Reconstitution

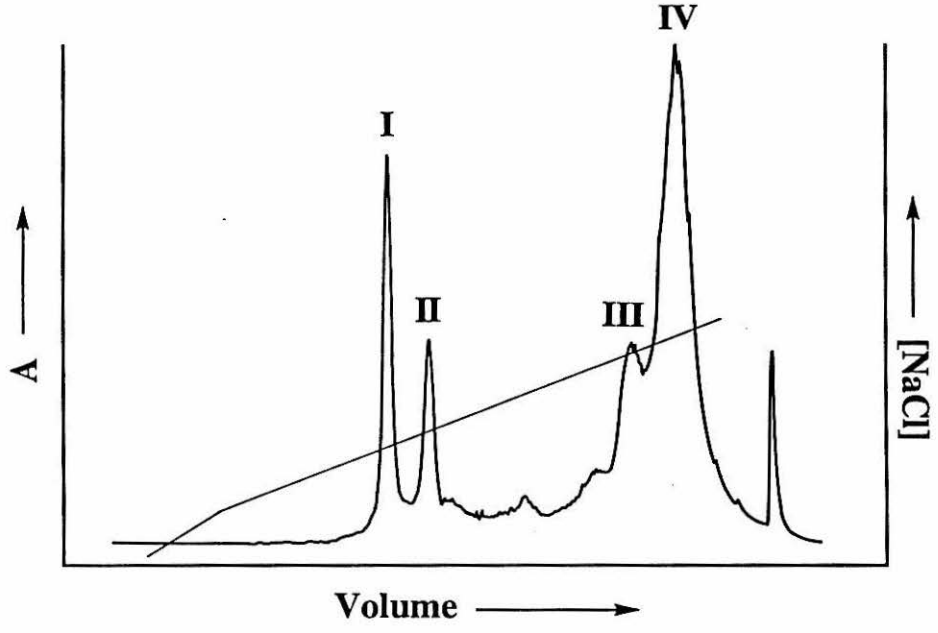
Peptide (10–20 mgs) was dissolved in a small volume of water and/or low ionic-strength NaP_i (pH 7.0). Occasionally, sonication was required to thoroughly dissolve the peptide. 0.8 equivalent of the peptide in solution was added to 1 equivalent of cyt *c*(1–65)lactone in a pear-shaped flask to give a final concentration of 0.3–0.4 mM cyt *c*(1–65)lactone. A dual manifold vacuum/argon Schlenk line was utilized to degas the solution by alternately purging and filling with argon which had been scrubbed with manganese oxide. At least 40 purge/fill cycles were performed, and care was taken to assure that the solution did not bubble excessively. The solution was reduced in a Vacuum Atmospheres HE-493 Dri-Train glove box by addition of 1.5–3 equivalents of a sodium dithionite (~ 2 mg/mL solution in water). The sodium dithionite was added to the reaction mixture dropwise until a color change could be observed, indicating reduction. A slight excess of dithionite (~ 0.5 equivalent) was added to assure that the solution remained reduced. The reaction was allowed to proceed with very gentle stirring for at least 40 hours. Occasionally, oxidation of the solution was observed during the course of the reaction, necessitating the addition of more dithionite. The semisynthetic protein was purified on a Mono S HR 10/10 column in 25 mM NaP_i (pH 7.0) with a gradient running from 0 to 38 mM NaCl over 100 mL (Figure 2.13).

Characterization of Cytochrome *c* Variants

Electrophoresis

SDS/PAGE gels (15%) were run at 10 mA using a Bio-Rad Model 500/200 power supply and stained with Sigma Brilliant Blue.

Figure 2.13 Typical cation-exchange (FPLC Mono S HR 10/10) chromatogram for the reconstitution of horse heart Ala80cyt *c*. Peak I is the reduced product and peak II is the oxidized product. The major peaks consist of unreacted cyt *c*(1–65) (peaks III and IV). The thin line represents the gradient of NaCl. Detection is at 280 nm on the 1.0 absorbance scale.


IV
I
II
III
A
[NaCl]
Volume

Capillary Zone Electrophoresis (CZE) was performed at the Caltech Biopolymer Synthesis and Analysis Resource Center by John Racs. Samples (~ 0.1 mg/mL protein) were provided in low ionic strength (12.5 mM) NaP_i (pH 7.0) and run at pH 2.5, at + 30 kV and a temperature of 40 °C with a sampling interval of 0.1 seconds.

Absorption Spectroscopy

UV/vis absorption spectra for routine use were taken on a Hewlett-Packard 8452 diode-array spectrometer run by a PC. High-resolution spectra were obtained on a Cary-14 spectrophotometer driven by On-Line Instruments Service (OLIS) software on a PC. All spectra were collected at room temperature in 1.000-cm pathlength quartz cuvettes. A cell buffer baseline was automatically subtracted from routine spectra. For data obtained on the Cary-14, a cell buffer baseline was collected and manually subtracted from the data using the OLIS software.

Extinction coefficients for mutants were determined by the pyridine hemochrome method as described by Berry and Trumpower.[30] The absorption spectrum of the fully oxidized protein from a stock solution was first recorded. The protein solution was then diluted with a stock solution of NaOH and pyridine to give a final concentration of 100 mM NaOH and 20% v/v pyridine. The absorption spectrum of the product was recorded immediately. The concentration of heme was then calculated using the extinction coefficient at 535 nm of 10.24 $mM^{-1}cm^{-1}$. This solution was reduced with excess dithionite and the absorption spectrum recorded. The concentration of heme was calculated using the extinction coefficient at 550 nm of 30.27 $mM^{-1}cm^{-1}$.[30]

Circular Dichroism Spectroscopy

Circular Dichroism (CD) spectra were collected on a Japan Spectrometer Company (JASCO)-600 CD spectropolarimeter in a jacketed 0.100-cm pathlength quartz cell. The sample was typically ~ 12 μM protein in 25 mM NaP_i (pH 7.0). Samples used

for variable temperature work or for measurements below 200 nm were degassed by the purge/fill method. Data were collected at 50 nm/min with a 1-nm bandwidth, 0.5-s time constant, 20-mdeg sensitivity, and a 0.2-nm step resolution. A cell buffer baseline was collected and subtracted from data. 4 or 8 scans were collected for each spectrum. Data were smoothed using a Fourier transform algorithm that removed high-frequency components. For variable temperature work, the temperature was controlled using a Brinkman Lauda K-2/R circulator and water bath, allowing the temperature to equilibrate for five minutes at each interval. The fraction of protein unfolded was calculated from the intensity of the CD signal at 222 nm using the equation:

$$\text{fraction unfolded} = (CD_T - CD_f)/(CD_{unf} - CD_f) \qquad (2.1)$$

where CD_T is the CD signal at temperature T, CD_f is the CD when the protein is fully folded (determined at 25 °C), and CD_{unf} is the CD when the protein is denatured. Protein unfolding curves were made by plotting the fraction of unfolded protein vs. temperature using the plotting program Kaleidagraph.

Electron Paramagnetic Resonance Spectroscopy

Electron paramagnetic resonance (EPR) data were obtained on a Bruker ESP 300 instrument with a liquid helium cryostat with the assistance of Prof. David G Goodin at The Scripps Laboratory and Research Institute in La Jolla, California. The 0.1 mM Fe(III)Ala80cyt *c* sample was in a 50% 100 mM HEPES (pH 7.0) 50% glycerol solution in a quartz EPR tube and was flash-frozen in liquid nitrogen to make a glass before inserting into the instrument cavity. The sample was run at 33 K and 1 mW power. Spectra of native Fe(III)cyt *c* were obtained on a 0.5 mM sample at 6.8 K in 50% 100 mM HEPES, pH 7.0, 50% glycerol at 80 μW power.

Ligand-field parameters were calculated from the measured **g** values as described by Taylor.[31] The equations used to calculate ligand-field parameters Δ and *V* (axial and rhombic distortion, respectively) are:[31,32]

$$a = (g_z + g_y)/4p \tag{2.2}$$

$$b = g_z - g_x/4p \tag{2.3}$$

$$c = (g_y - g_x)/4p \tag{2.4}$$

$$2p^2 = g_z + g_y - g_x \tag{2.5}$$

$$E/\lambda = -(b + c)/2a \tag{2.6}$$

$$A = E/\lambda + (a + c)/2b \tag{2.7}$$

$$B = E/\lambda + (a + b)/2c \tag{2.8}$$

$$V/\lambda = A \tag{2.9}$$

$$\Delta/\lambda = B - A/2 \tag{2.10}$$

where $|\Delta/\lambda|$ is the tetragonal field, $|V/\lambda|$ is the rhombahedral field, and $|V/\Delta|$ is the rhombicity. The **g** values used were those of the improper axis system, with

$$g_z g_y g_x > 0. \tag{2.11}$$

To transform the **g** values to the improper axis system, the following substitutions were made, with the measured **g** values on the left hand side:[33]

$$g_z \rightarrow -g_x \tag{2.12}$$

$$g_y \rightarrow g_z \tag{2.13}$$

$$g_x \rightarrow -g_y \tag{2.14}$$

Magnetic Circular Dichroism Spectroscopy

Dr. Barton Hawkins carried out magnetic circular dichroism (MCD) measurements in the laboratories of Prof. John Dawson at the University of South Carolina. MCD spectra were obtained on a JASCO J-500A spectropolarimeter equipped with a JASCO MCD-1B electromagnet operated at a field strength of 1.41 T. The J-500A was interfaced to an IBM PS/2 Model 50 by a JASCO IF-500-2 interface unit. 2–4 scans were signal-averaged per spectrum. Data were collected at 4 °C on 8.9–45 μM samples in 25 mM NaP_i (pH 7.0) or in 25 mM NaOAc (pH 4.5). Data handling was as previously described.[34]

Nuclear Magnetic Resonance Spectroscopy

500 MHz ^{1}H NMR spectra were collected on a Bruker AMX 500 spectrometer. Samples were 0.25–1 mM in protein in 25 mM NaP_i in D_2O. The pH of NMR samples was adjusted by addition of small volumes of 2M NaOH or 1M HCl. Reported pH values are not corrected for the isotope effect. The residual HDO peak was suppressed by presaturation.

Ligand-Binding Studies

Preparation of Ligand-Bound Derivatives

Derivatives of ferrous proteins were prepared from reduced, deoxygenated samples which were made as follows. A purified solution of protein was degassed on a Schlenk line by the purge/fill method (~ 50 cycles). A slurry of Sephadex G-25 and 25 mM NaP_i (pH 7.0) buffer was also degassed by the purge/fill method. 25 mM NaP_i (pH 7.0, ~ 100 mL) was degassed on a Schlenk line by the freeze/pump/thaw method. The protein sample was reduced with a small (5–10x) excess of solid sodium dithionite in the glove box, immediately loaded on a column (20 cm x 1 cm) packed with the Sephadex, and eluted with the NaP_i buffer. The reduced protein was collected in a laser cuvette and sealed. To produce the CO and O_2 derivatives, 1 atm of CO or O_2 was introduced to the cell through a septum.

The O_2 derivatives of Ala80cyt *c* and of Phe67Ala80cyt *c* are the major products purified from the reconstitution reactions. A quick way to prepare the O_2 derivative of a protein that is partially or fully oxidized is to reduce the sample with a 10x excess of solid sodium dithionite and immediately apply it to a PD-10 desalting column (Pharmacia), eluted with 50 mM NaP_i (pH 7.0) under aerobic conditions. O_2 derivatives were found to be more stable to autoxidation when kept under 1 atm of O_2.

Cyanide derivatives were prepared by adding a ~ 100x excess of NaCN to a sample of oxidized protein which was previously separated from reduced protein by

FPLC, or that was oxidized chemically with K[Co(EDTA)] or with $K_3[Fe(CN)_6]$. Excess ligand was removed by passage down a PD-10 column eluted with 25 mM NaP_i (pH 7.0). H_2O-Fe(III)Ala80cyt *c* was prepared by lowering the pH of the solution of Fe(III)Ala80cyt *c* below 6. Above pH 7, the Fe(III)Ala80cyt *c* has His-OH^- ligation. All of the ferric derivatives were found to be stable over time. The CN^- derivatives are particularly stable, with UV/vis and, in the case of CN-Ala80cyt *c*, NMR spectra that did not change over a period of many months.

Ligand Photolysis: Data Collection

Transient absorption experiments were performed at the Beckman Institute Laser Resource Center under the direction of Dr. Jay R. Winkler. The experiments were driven by a 386 PC with software written by Dr. Winkler and Dr. Max Bachrach. The excitation source for transient absorption experiments with an excitation wavelength of 480 nm was a Lambda Physik FL 3002 dye laser (coumarin 480 dye) pumped by a Lambda Physik LPX210 excimer laser. The excimer pump beam (300 nm, ~ 300 mJ) was attenuated with a beamsplitter (75% transmittance) before entering the dye laser. The 480-nm pulses from the dye laser had a pulse width (full width at half maximum) of 25 ns. The source of 532-nm excitation pulses was a frequency-doubled Q-switched Quanta-Ray Nd:YAG laser with a pulse width of 20 ns. The output of the excitation source laser was passed through a beam expander and through two polarizers. Laser pulse energy was measured to be between 1–3 mJ for the dye laser and 5–10 mJ for the Nd:YAG laser. A small portion of the pulse was directed onto an Electro-Optics Technology ET2000 pin photodiode which led to a Phillips Scientific Model 6930 windowing discriminator. Only pulses that fell in a specified energy range would trigger the digitizer for data collection.

A PTI 75-W xenon arc lamp in a PTI A1010 (f/15 ellipsoidal reflector) housing provided the probe light. The arc lamp power source was a PTI LPS-220 arc lamp supply. The probe light was passed through a 1-mm aperture and collected with a 50-cm

radius-of-curvature (ROC) spherical-concave mirror and focused on the sample. Light transmitted through the sample was collected by a second 50-cm ROC mirror and then focused onto the slit (1 mm) of an Instruments SA monochrometer. The excitation beam was directed through 1-cm holes in the center of the 50-cm ROC mirrors. The pump and probe lights were aligned such that the probed region of the sample was fully within the laser beam profile. Transient absorption signals were detected using a Products for Research R928 photomultiplier tube (PMT) housed in a thermoelectrically cooled chamber. The transient signal was amplified using a quasi-differential amplifier (200 MHz, 7 μA/mV) for timescales between 1.25 and 750 μs, or a LeCroy DSP 1402E amplifier for timescales between 1 and 1000 ms. The signal was digitized using a Sony Tektronix RTD digitizer and data were transferred to a PC.

Transient absorption experiments were performed at room temperature on 3–6 μM protein in 25 mM NaP_i (pH 7.0) under air, 1 atm O_2, or a known volume of CO. To introduce the CO, a flask fitted with a septum was flushed with 1 atm of CO. A Hamilton air-tight syringe was then used to introduce a small volume of argon into the flask of CO to prevent back-flow of atmospheric O_2 upon removal of the syringe; the volume of CO was then drawn from the flask and added to the protein sample through a septum. Samples were contained in 1 cm precision quartz cuvettes (NSG Glassware) connected *via* a graded seal to a glass bulb and a Kontes 24/40 joint for degassing. Samples were stirred continuously with a Teflon stirbar during data collection on short timescales and only between laser shots on long timescales. On short timescales, a laser repetition rate of 5 Hz was used, while on long timescales that required use of the slow amplifier a repetition rate of 0.7 Hz was used to reduce 60-Hz interference. Appropriate optical interference filters were placed before the sample. Neutral density filters and PMT voltage were adjusted to maximize signal-to-noise ratio and to effectively utilize the PMT dynamic range. Ligand recombination was monitored near 430 nm (all samples) and near

414 nm (CO) or 406 nm (O_2). The instrument response was obtained by collecting data in the luminescence decay mode at 480 nm.

Ligand Photolysis: Calculations

The rate constants given by transient absorption data were determined by fitting the data with least-squares fitting programs written by Dr. Jay Winkler. Most data were fit using KINFIT which allows the selection of the portion of the data set to be fit and the time zero point. The data can be fit using up to three exponential terms according to the equation

$$y(t) = c_0 + c_1 \exp(-k_1 t) + c_2 \exp(-k_2 t) + c_3 \exp(-k_3 t) \quad (2.15)$$

where the k_n are first-order rate constants and the c_n constants. Data on short (< 2.5 ms) timescales were fit by using the program DECON which deconvolutes the measured instrument response from a data file while fitting up to three exponentials using a least-squares procedure like that of KINFIT.

Laser photolysis experiments were performed under conditions of pseudo-first-order kinetics so that for the reaction

$$\text{Protein} + \text{L} \xrightarrow[k_{on}]{} \text{Protein–L} \quad (2.16)$$

[Protein] << [L], giving a pseudo-first-order rate law for ligand recombination of

$$d[\text{Protein-L}]/dt = k_{ps}[\text{Protein}] \quad (2.17)$$

where

$$k_{ps} = k_{on}[\text{CO}]. \quad (2.18)$$

In this case, the observed rate constant given by a first-order fit at a given [CO] is k_{ps}. The concentration of CO in solution was found through the relationship of its solubility and its partial pressure over aqueous solution:[35]

$$[\text{CO}] = 1.36\ \mu\text{M/mmHg} \quad (2.19)$$

The reported ligand-binding rates were obtained by error-weighted averaging of the relevant fits.[36] The error in the rate obtained from each fit was estimated by comparison of the quality of fits to a range of rate values.

The quantum yield Φ_{rel} of photodissociation was calculated relative to that observed for the CO derivative of h. h. Fe(II)Mb for which the quantum yield for photodissociation of CO from the protein matrix is known to be close to 1.[12] This method simplifies the determination of the quantum yield as it does not require explicit consideration of the laser power dependence of photodissociation. The protein samples for photodissociation experiments were at concentrations such that the absorbance at the excitation wavelength was nearly the same for the cytochrome and Mb samples. Transient absorption data on CO-Fe(II)cytochrome variants and CO-Fe(II)Mb were collected under conditions of the same excitation wavelength and the same laser power. The relative quantum yield for photodissocation of the cyt *c* mutant on the observed timescale is then given by

$$\Phi_{rel} = \phi_{cyt}/\phi_{Mb} \tag{2.20}$$

where ϕ for the cyt *c* variant or for Mb is given by

$$\phi = C_{tr}/C_{tot} \tag{2.21}$$

where C_{tr} is the maximum transient concentration of Fe(II) protein after photodissociation and C_{tot} is the total concentration of the protein in the sample. This quantity can be related to the transient absorption data through the extinction coefficients for the deoxy-Fe(II)and the CO-Fe(II) derivatives of either the cyt *c* variant or Mb, ε_{deoxy} and ε_{CO}, respectively. Multiplying equation (2.21) by $\Delta\varepsilon/\Delta\varepsilon$ where $\Delta\varepsilon$ is defined as

$$\Delta\varepsilon = \varepsilon_{deoxy} - \varepsilon_{CO} \tag{2.22}$$

gives

$$\phi = \Delta\varepsilon\, C_{tr}/\Delta\varepsilon\, C_{tot} \tag{2.23}$$

$$\phi = \Delta A_{tr}/\Delta A_{ss} \tag{2.24}$$

where ΔA_{tr} is the maximum observed optical density change in the transient absorption experiment on the timescale observed and ΔA_{ss} is the change in absorption between the fully reduced and deoxygenated protein and the CO derivative as calculated from the known protein concentration and extinction coefficients.

The observed quantum yield of ligand photodissociation for ferrous heme proteins typically depends strongly on the timescale observed. Here, quantum yield is defined as the yield of fully dissociated product capable of rebinding in bimolecular reactions as is observed on timescales in the millisecond range.

Determination of Autoxidation Rates

The autoxidation rate k is the rate constant for the reaction

$$\text{Protein(II)-O}_2 \xrightarrow[k]{-O_2^-,\,+L} \text{Protein(III)-L} \qquad (2.25)$$

where L may be H_2O, OH^-, or nonexistent. The reaction was initiated by exposing a reduced sample to air. A UV/vis spectrum of the sample was taken immediately after preparation. The sample was maintained in a sealed cuvette to prevent evaporation and kept at room temperature which was observed to be 22 ± 1 °C. After 250–350 hours, the autoxidation had progressed significantly and the sample was oxidized by the addition of a 100x excess of $K_3[Fe(CN)_6]$ and the final spectrum taken. A time of about 3 hours was required to assure that oxidation by $K_3[Fe(CN)_6]$ was complete. A plot of $(A_\infty - A_t)$ vs. time (in hours) was made, where A_t is the absorption at wavelength λ at time t and A_∞ is the absorption at λ in the spectrum of the fully oxidized protein. This plot was fit to the first-order equation

$$(A_\infty - A_t) = C * \exp(-kt) \qquad (2.26)$$

where C is a constant and k is the autoxidation rate in units of h^{-1}.

Determination of Dioxygen Affinity

The affinity of Fe(II)Ala80cyt *c* for O_2 was determined by observing the change in the UV/vis spectrum of a sample of fully deoxygenated and reduced Fe(II)Ala80cyt *c* (prepared as described above) upon the introduction of small amounts of dioxygen. The sample was maintained in a sealed cell of known volume with a degassing bulb. The cell, filled with argon, was fitted with a Kontes stopcock and a Sub-a-Seal septum. Small (μL) amounts of 100% dioxygen were introduced with a Hamilton gas-tight syringe, and the UV/vis spectrum monitored. It was found that approximately one hour was needed for equilibration after each addition of dioxygen. In spite of the slow autoxidation rate exhibited by O_2-Fe(II)Ala80cyt *c* (*vide infra*), at low dioxygen pressures autoxidation was fast enough to prevent the collection of many data points.

To calculate the binding constant for O_2, the mole-fraction of dioxygen-ligated protein present at each oxygen pressure, Y_i, was determined from the absorption spectra using the equation:

$$Y_i = [A_{obs} - A_p]/[A_{pO_2} - A_p] \tag{2.27}$$

where A_{obs} is the observed absorbance at a given wavelength λ, A_{pO_2} is the absorption of the fully oxygenated protein at λ, and A_p is the absorption of the fully deoxygenated and reduced protein at λ. The partial pressure of dioxygen in the cuvette after each addition of dioxygen was calculated from the equation:[35]

$$P_i = [(P_0\, V_i)/(V_c - V_s)] - (nY_iRT)/(V_c - V_s) \tag{2.28}$$

where $P_0 = 760$ mmHg, V_i = total volume of O_2 added, P_i = partial pressure of O_2 in the cuvette at volume V_i, V_c = total volume of cuvette, V_s = volume of sample, n = moles of heme in the sample, and Y_i = mole faction of oxygen-ligated species. The fraction of oxygen-bound protein was then plotted against the partial pressure of oxygen and the resulting data fit to the Hill equation assuming the Hill coefficient is 1:[12]

$$Y_i = P_{O_2}/[P_{O_2} + P_{1/2}] \tag{2.29}$$

where P_{O_2} is the partial pressure of O_2 and $P_{1/2}$ is the partial pressure of oxygen at which one half of the protein has dioxygen bound. $P_{1/2}$ is related to the O_2 association constant K according to[35]

$$K = [P_{1/2}(\text{mmHg}) * 1.82\ \mu\text{M/mmHg}]^{-1} \quad (2.30)$$

Molecular Modeling

The 1.94-Å oxidized h. h. cyt *c* structure[3a] and the 1.6-Å sperm whale Mb structure were used in molecular modeling.[13] Modeling was performed using Biograf version 2.2 (BioDesign, Inc.) on a VaxStation 3500 equipped with a PS340 Evans and Sutherland graphics system. The Dreiding force field was utilized for energy minimizations.

Results and Discussion

Characterization of Horse Heart Cytochrome *c* (1-65)

The CNBr cleavage reaction routinely gave a 50–60% yield of pure cyt *c*(1–65) lactone product. Cyt *c*(1–65) was characterized by SDS/PAGE electrophoresis, absorption spectroscopy, and CD spectroscopy. SDS/PAGE indicated the expected molecular weight of ~ 8000, with formation of only a small amount of dimerized product. As previously reported, the absorption spectrum of Fe(III)cyt *c*(1–65) (Figure 2.14) is similar to that of native Fe(III)cyt *c* in the visible region, with absorption maxima of 407 and 525 nm.[6a] As anticipated, bands in the UV portion of the spectrum due to peptide backbone and aromatic side chains have less intensity than native cyt *c*, consistent with the absence of residues 66–104. In addition, the 695-nm band (the Met80 → Fe(III) charge transfer) is absent in the cyt *c*(1–65) spectrum. The CD spectrum of the 1-65 fragment indicates that it has no α-helical content (Figure 2.15).

Figure 2.14 UV/vis absorption spectrum of 8 μM cyt *c*(1–65) in 25 mM NaP_i (pH 7.0). The spectrum was obtained at room temperature in a 1-cm cuvette.

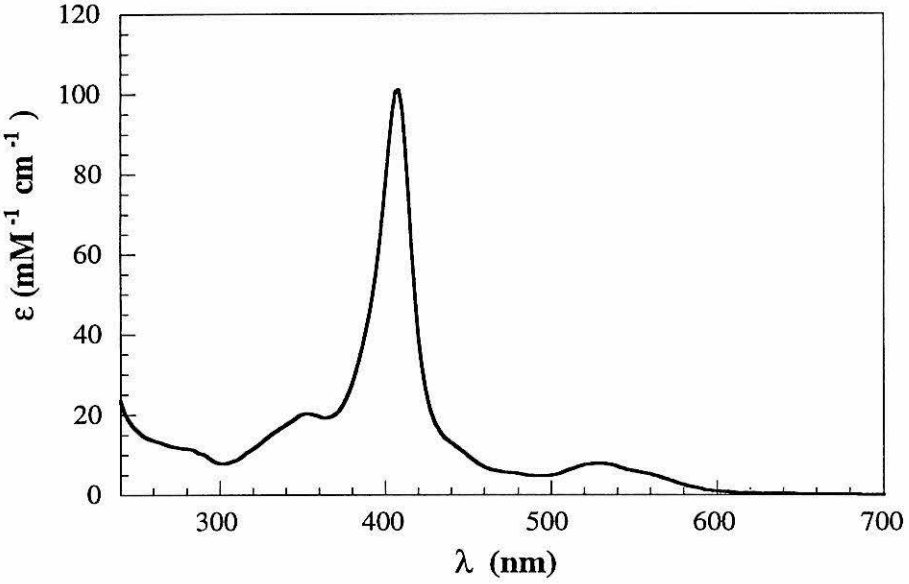
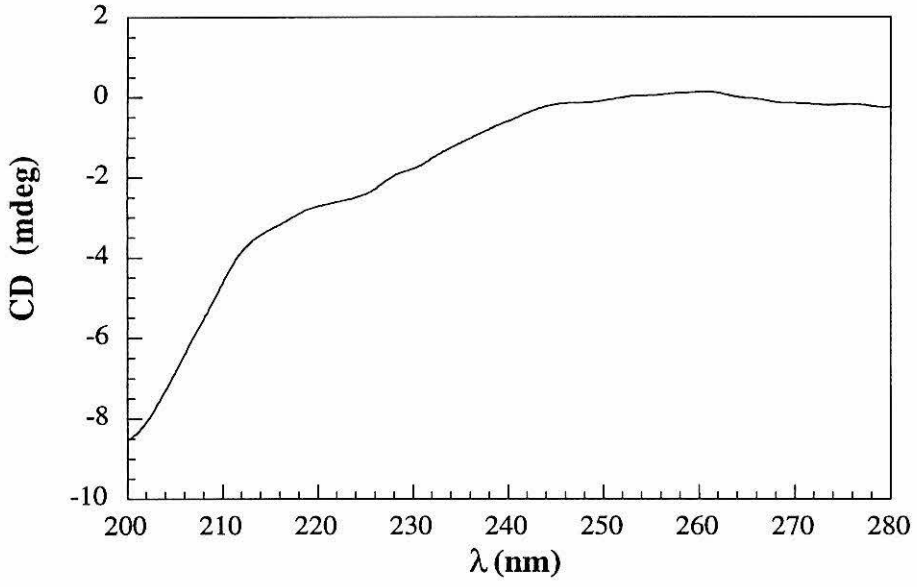

120
100
80
60
40
20
0
ε (mM-1 cm-1)
300
400
500
600
700
λ (nm)

Figure 2.15 Far-UV CD spectrum of 12 μM cyt *c*(1-65) in 25 mM NaP_i (pH 7.0) at room temperature in a 0.1-cm cell. The spectrum is characteristic of an unfolded polypeptide chain.


2
0
-2
-4
-6
-8
-10
CD (mdeg)
200
210
220
230
240
250
260
270
280
λ (nm)

Characterization of Semisynthetic Ala80cytochrome *c*

Chromatography and Electrophoresis

Cation exchange chromatography of the Ala80cyt *c* reconstitution reaction mixture results in facile separation of reactants from products (Figure 2.13). The product is similar, but not identical, to native protein in its chromatographic elution profile. The difference from native cyt *c* may be due to minor structural changes induced by the Met80Ala mutation or to changes in the effective charge of the protein due to the change in heme ligation. The elution times of both native and Ala80cyt *c* have been observed to be very sensitive to heme oxidation state. For Ala80cyt *c*, the fractions constituting the first peak to elute (peak I, 21 mM NaCl) were pink/orange in color. The fractions in peak II (23 mM NaCl) were red/brown. It was initially verified that the two forms of the protein represent oxidized and reduced forms of the product by chemical oxidation and reduction. Peak I (Fe(II)Ala80cyt *c*), when oxidized with $K_3[Fe(CN)_6]$, yields a product with spectroscopic and chromatographic properties indistinguishable from those of peak II. Peak II (Fe(III)Ala80cyt *c*) takes on a spectrum and chromatographic elution time identical to those of peak I upon reduction with sodium dithionite followed by gel filtration under aerobic conditions. In the FPLC purification, the oxidized and reduced forms of Ala80cyt *c* are well separated from each other and also from the unreacted cyt *c*(1–65) lactone and inactive cyt *c*(1–65) carboxylate. Typical yields of Ala80cyt *c* are 5%–15%, compared with a 70% yield for the reconstitution of cyt *c*(1–65) with a peptide containing Met80.[9e] This result supports previous observations that the presence of an axial ligand at position 80 in the peptide is important but not essential for the formation of a productive complex between cyt *c*(1–65) and the 66–104 peptide.[9cdg] Hydrophobic interactions between the heme and the 66–104 peptide likely provide the driving force for reconstitution of Ala80cyt *c*.

Ala80cyt *c* was found to be similar to native cyt *c* in its electrophoretic properties. An SDS/PAGE experiment in which the mobility of Ala80cyt *c* under denaturing

conditions was compared to that of native cyt *c* and cyt *c*(1–65) demonstrated that Ala80cyt *c* is homogeneous and has a molecular weight very close to that of the native protein. CZE analysis gives sharp, single peaks with elution times of 7.8 min. (Fe(II)Ala80cyt *c*) and 8.0 min. (Fe(III)Ala80cyt *c*) which are the same as those for native Fe(II)cyt *c* and Fe(III)cyt *c*, confirming that a homogeneous cyt *c*-like protein was formed.

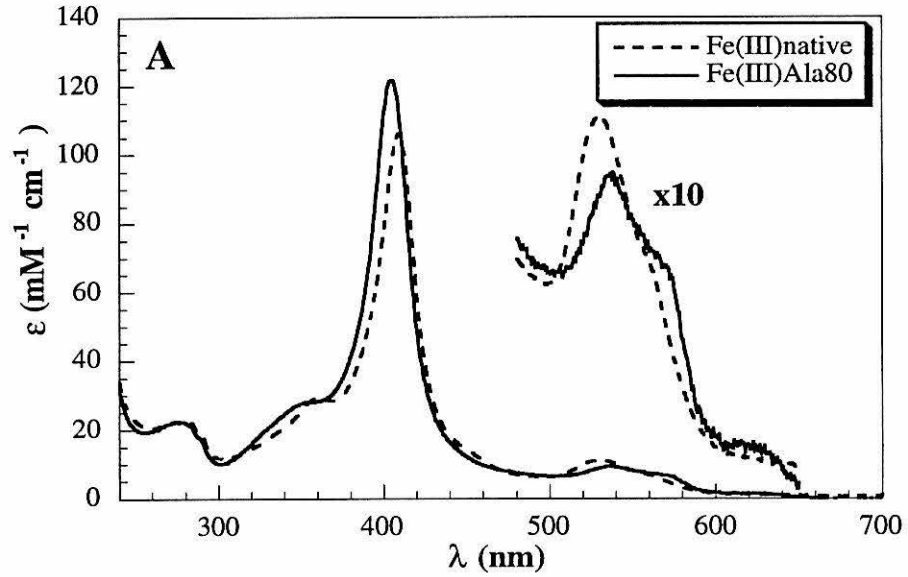
Absorption Spectroscopy

The absorption spectra of Fe(III)Ala80 and Fe(II)Ala80 were found to be dramatically different from those of the native protein (Figure 2.16). Compared to native h. h. Fe(II)cyt *c*, h. h. Fe(II)Ala80 as purified has a Soret band blue-shifted by 8 nm to 406 nm and α, β bands at 573 and 538 nm, compared to 550 and 520 for native. The α, β bands of the reduced mutant have an intensity and shape similar to each other, while native cyt *c* has an α band which is more intense and sharp than its β band. Semisynthetic Fe(III)Ala80 has a Soret band red-shifted from native by 5 nm to 405 nm and three poorly-defined bands in the α, β region at 622, 569, and 538 nm. These significant spectroscopic differences from native cyt *c* were expected; it has been observed previously that changes in heme axial ligands result in significant perturbations to spectroscopic properties.[9cd] The extinction coefficients of the mutant proteins were also observed to change from those of native. The accuracy of the pyridine hemochrome method of extinction coefficient determination was verified through control experiments on native protein.

Circular Dichroism Spectroscopy

Far-UV CD spectroscopy was used as a probe of the secondary structure and stability of the Ala80 variant. The CD spectra of Fe(III)Ala80 and native Fe(III)cyt *c* at ambient temperature in the 190–280 nm region were found to be virtually

Figure 2.16 Superposition of the UV/vis absorption spectra of Ala80cyt *c* (solid lines) as purified from the reconstitution reaction and native h. h. cyt *c* (dashed lines) in the (A) oxidized and (B) reduced states. The reduced form of Ala80cyt *c* has dioxygen bound to the heme, and the oxidized form has been identified to have hydroxide ligation. Spectra were recorded on 8 μM protein samples in 25 mM NaP_i (pH 7.0) at room temperature in a 1-cm cell.


A
Fe(III)native
Fe(III)Ala80
x10
ε (mM-1 cm-1)
λ (nm)
0
20
40
60
80
100
120
140
300
400
500
600
700

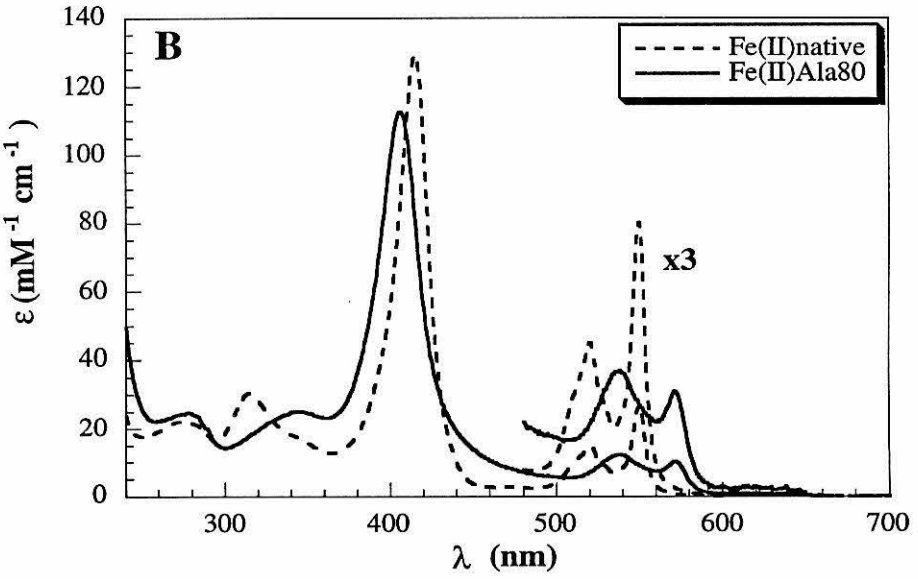

B
Fe(II)native
Fe(II)Ala80
x3
ε (mM-1 cm-1)
λ (nm)
0
20
40
60
80
100
120
140
300
400
500
600
700

Figure 2.17 Far-UV CD spectra of h. h. Fe(III)Ala80cyt *c* (solid line) and native h. h. Fe(III)cyt *c* (dashed line). Samples were 12 μM in protein in 25 mM NaP_i (pH 7.0) in a 0.1-cm cell. Spectra were recorded at ambient temperature.

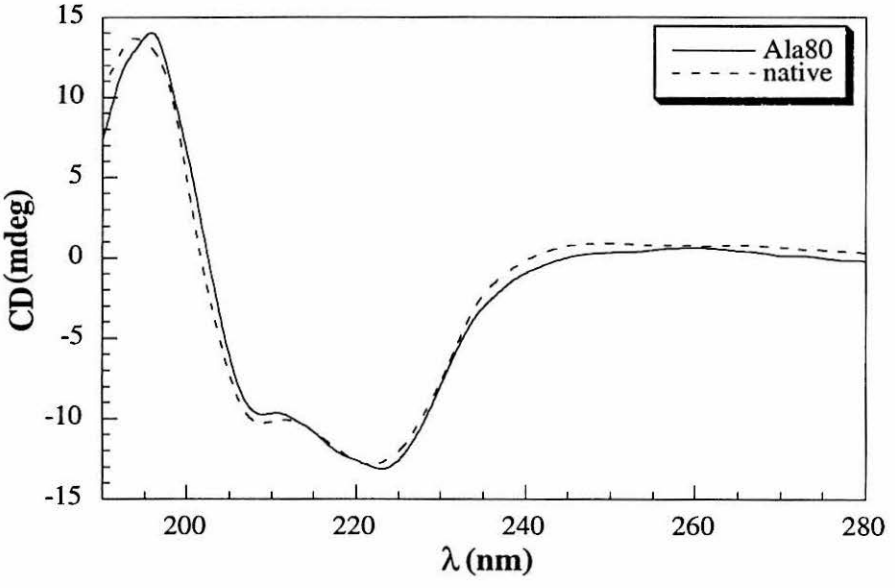
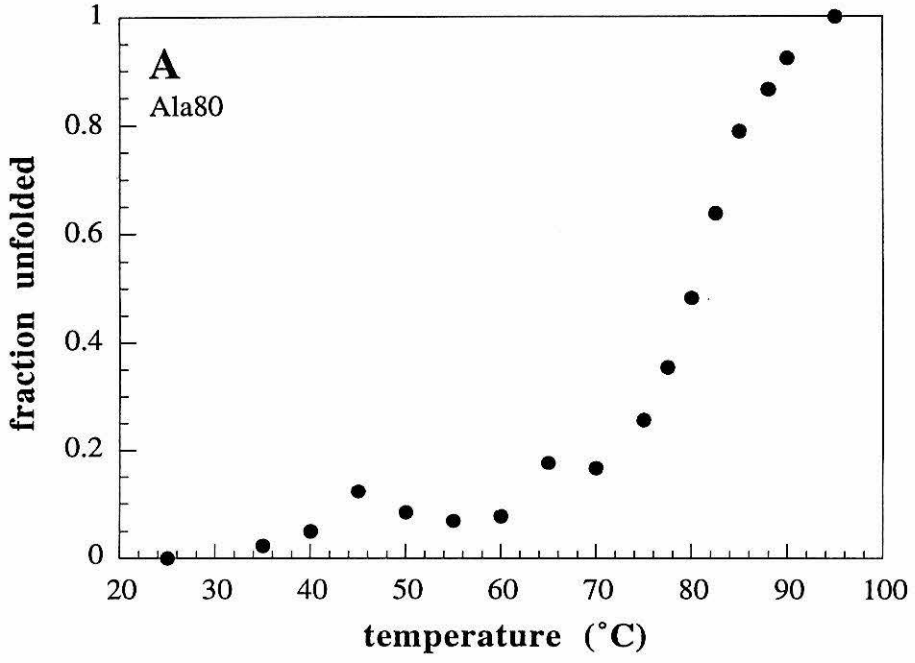
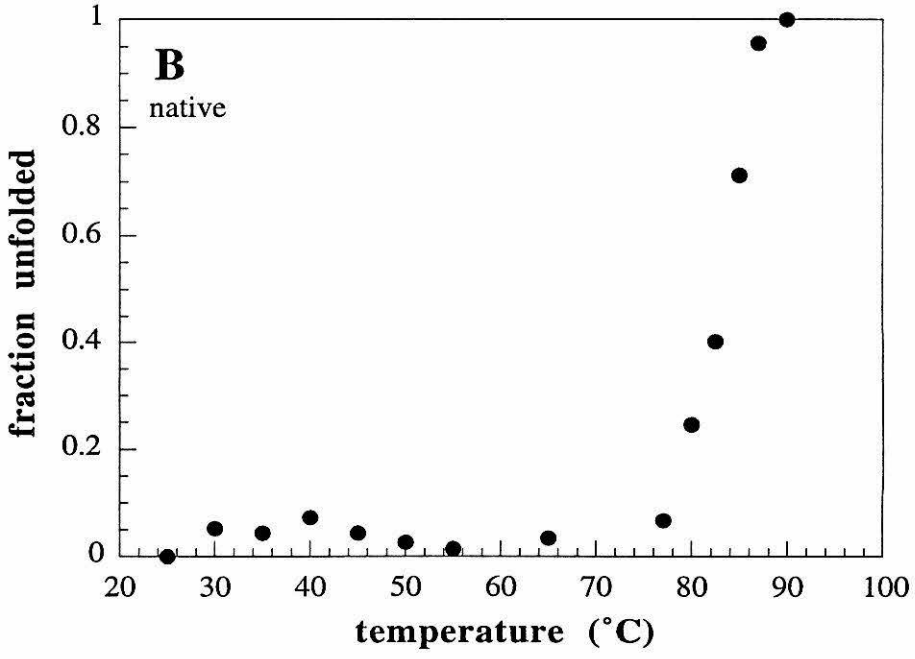

15
10
5
0
-5
-10
-15
CD(mdeg)
200
220
240
260
280
λ (nm)
Ala80
native

Figure 2.18 CD melting curves for (A) Fe(III)Ala80cyt *c* and (B) native Fe(III)cyt *c*. Samples were 10 μM in protein in 25 mM NaP_i (pH 7.0) in a 0.1-cm jacketed cell. The transition temperatures for the protein are T_M = 85 °C for native cyt *c* and T_M = 81 °C for Ala80cyt *c*.


A
Ala80
fraction unfolded
temperature (°C)
0
0.2
0.4
0.6
0.8
1
20
30
40
50
60
70
80
90
100


B
native
fraction unfolded
temperature (°C)
0
0.2
0.4
0.6
0.8
1
20
30
40
50
60
70
80
90
100

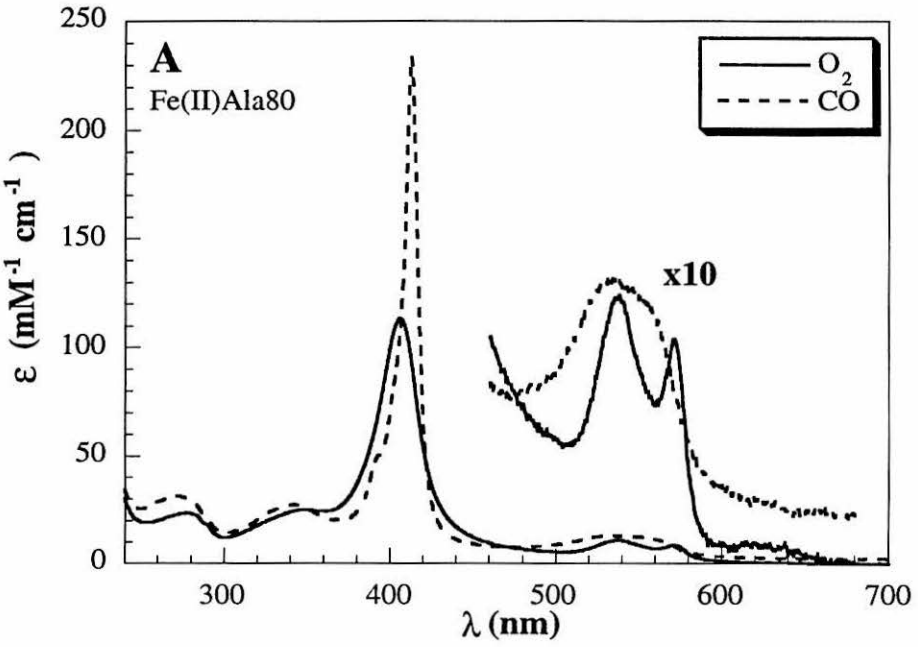
superimposable, as shown in Figure 2.17. These spectra are typical for α-helical proteins, with a maximum at 195 nm and minima at 205 and 222 nm. To determine the thermal stability of the mutant, the change in CD at 222 nm was monitored as temperature was raised (Figure 2.18). The CD spectra roughly maintained isosbestic points, supporting the two-state transition theory. CD baseline drift, which can be considerable over the course of an experiment, accounts for much of the loss of isosbestic points. The midpoints for the melting transitions was found to be 81 °C for h. h. Fe(III)Ala80 compared to 85 °C for native h. h. Fe(III)cyt *c*, indicating that the Met80Ala mutation slightly decreases the thermal stability of the cyt *c* fold. This difference may be due to the loss of the Met-Fe bond or to indirect effects such as destabilization of the hydrophobic core, disruption of the hydrogen-bond network involving Met80 in native cyt *c*, or minor structural changes near and distant from the mutation site.

Ligand Binding to Horse Heart Ala80cyt *c*

Absorption Spectroscopy

It was anticipated that Ala80cyt *c* would have spectroscopic properties analogous to ligand-binding heme proteins such as Mb and peroxidases. The UV/vis absorption spectra of h. h. Ala80 derivatives compared to analogous Mb derivatives are given in Figures 2.19 and 2.20 (Fe(II) derivatives); 2.21 and 2.22 (Fe(III) derivatives). The absorption spectra maxima and extinction coefficients are presented in Table 2.2. In the reduced state under aerobic conditions, h. h. Ala80cyt *c* has an absorption spectrum remarkably similar to that of the dioxygen-bound form of h. h. Mb, providing an indication that dioxygen also binds Ala80cyt *c* (Figure 2.19). However, the absorption spectrum of deoxy-Fe(II)Ala80cyt *c* differs markedly from that of deoxy-Fe(II)Mb (Figure 2.20). Deoxy-Fe(II)Mb has a Soret band with a maximum at 434 nm as is characteristic for high-spin ferrous hemes.[2] Deoxy-Fe(II)Ala80 has a Soret band that appears to have two components: a major component with a maximum at 411 nm, and a

Figure 2.19 UV/vis absorption spectra of the O_2 (solid lines) and CO (dashed lines) adducts of (A) semisynthetic Fe(II)Ala80cyt *c* and (B) h. h. Fe(II)Mb. Samples were 6–10 μM in protein in 25 mM NaP_i (pH 7.0) in a sealed 1-cm cell.


A
Fe(II)Ala80
O2
CO
x10
ε (mM-1 cm-1)
λ (nm)
0
50
100
150
200
250
300
400
500
600
700

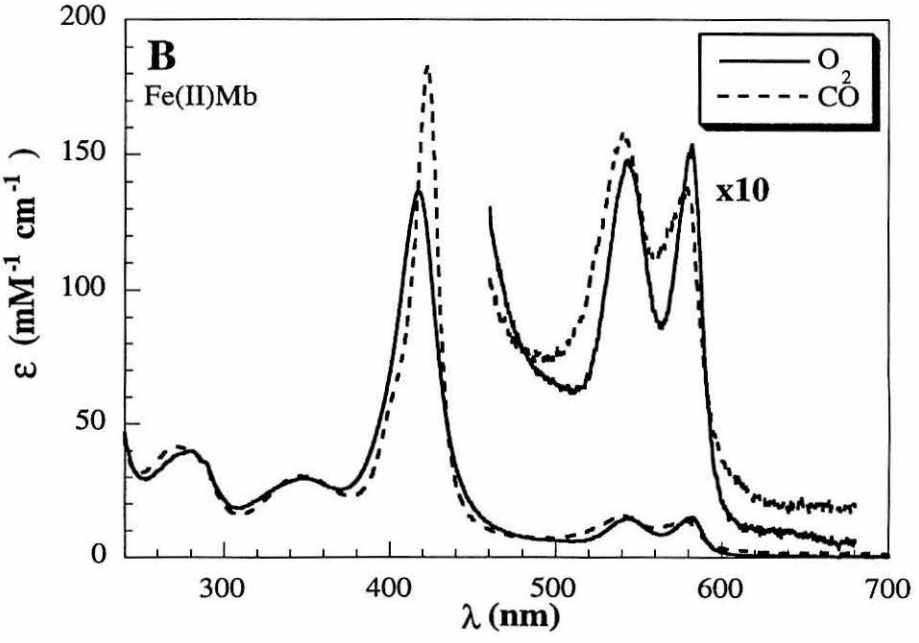
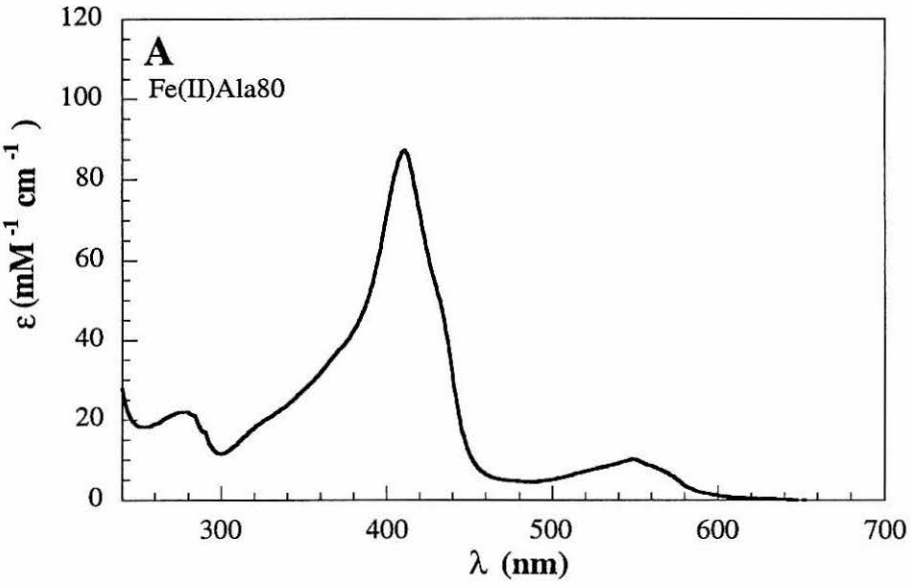

B
Fe(II)Mb
O2
CO
x10
ε (mM-1 cm-1)
λ (nm)
0
50
100
150
200
300
400
500
600
700

Figure 2.20 UV/vis absorption spectra of (A) h. h. Fe(II)Ala80cyt *c* and (B) h. h. Fe(II)Mb under anaerobic conditions. Samples are 8 μM in protein in 25 mM NaP_i (pH 7.0) in a sealed 1-cm cell.


A
Fe(II)Ala80
ε (mM-1 cm-1)
λ (nm)
0
20
40
60
80
100
120
300
400
500
600
700

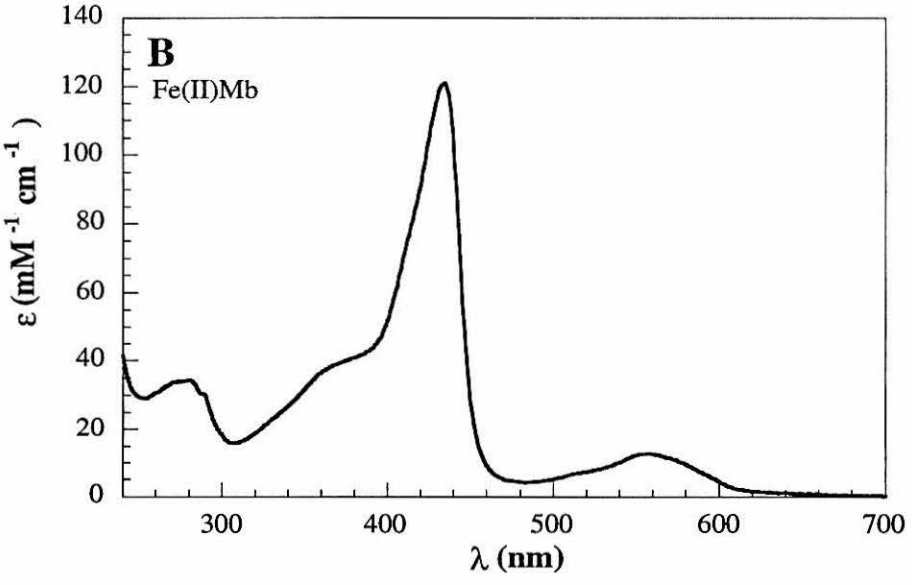
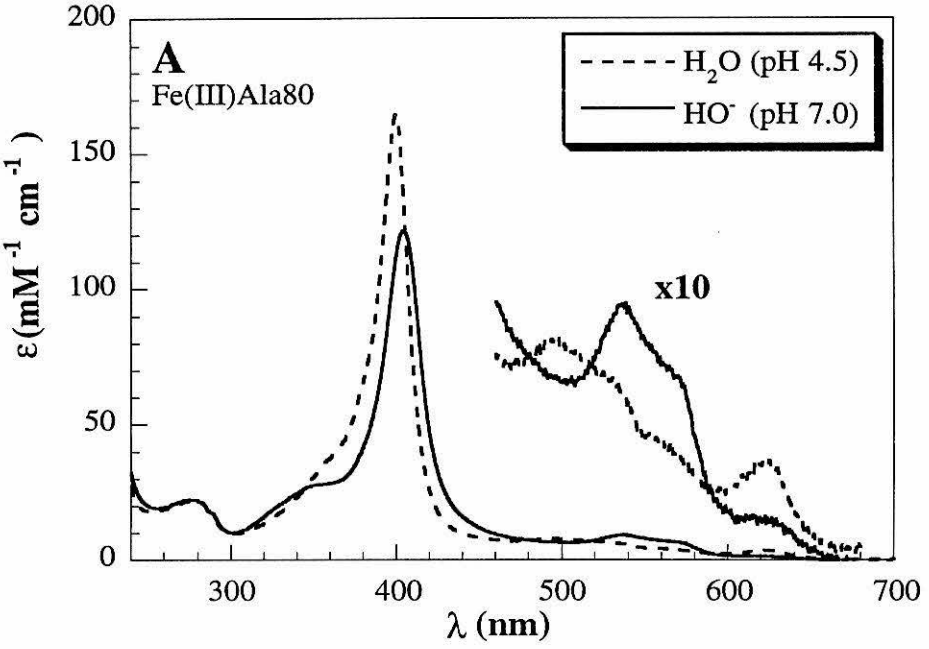

B
Fe(II)Mb
ε (mM-1 cm-1)
λ (nm)
0
20
40
60
80
100
120
140
300
400
500
600
700

Figure 2.21 UV/vis absorption spectra of the HO^- (solid lines) and H_2O (dashed lines) adducts of (A) semisynthetic h. h. Fe(III)Ala80cyt *c* and (B) h. h. Fe(III)Mb. Spectra were recorded at room temperature in a 1-cm cell. Samples were 6 mM in 25 mM NaOAc (pH 4.5), 25 mM NaP_i (pH 7.0), or 25 mM $NaHCO_3$ (pH 10.0).


A
Fe(III)Ala80
H2O (pH 4.5)
HO- (pH 7.0)
x10
ε(mM-1 cm-1)
λ (nm)
200
150
100
50
0
300
400
500
600
700

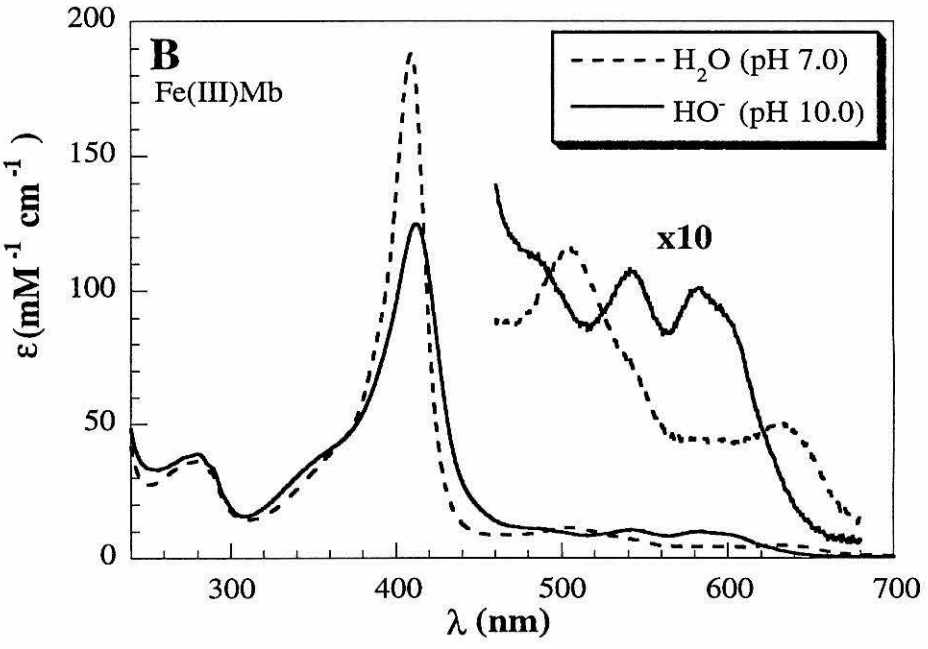
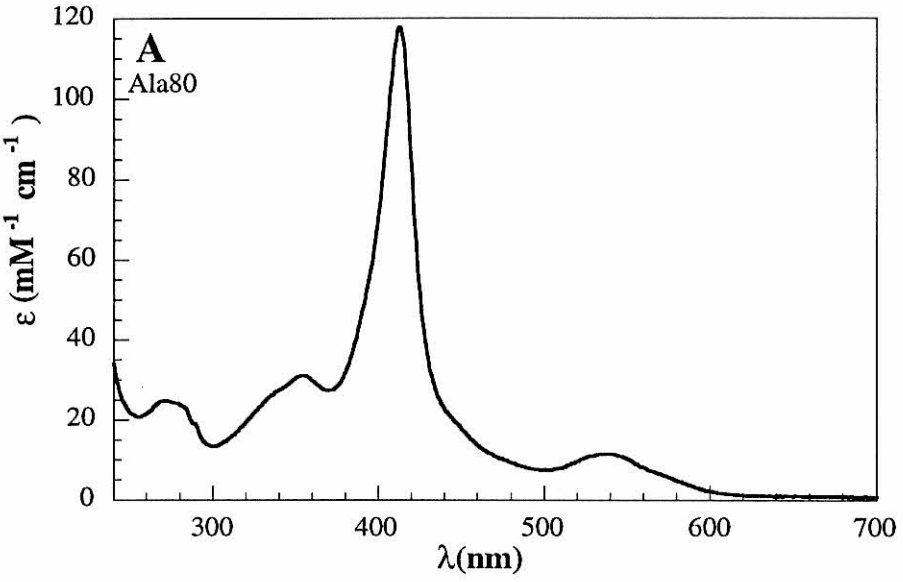

B
Fe(III)Mb
H2O (pH 7.0)
HO- (pH 10.0)
x10
ε(mM-1 cm-1)
λ (nm)
200
150
100
50
0
300
400
500
600
700

Figure 2.22 UV/vis absorption spectra of the cyanide adducts of (A) h. h. Fe(III)Ala80cyt *c* and (B) h. h. Fe(III)Mb. Spectra were taken at room temperature on 6 μM protein samples in 25 mM NaP_i (pH 7.0) in a 1-cm cell.


A
Ala80
ε (mM-1 cm-1)
λ(nm)
0
20
40
60
80
100
120
300
400
500
600
700

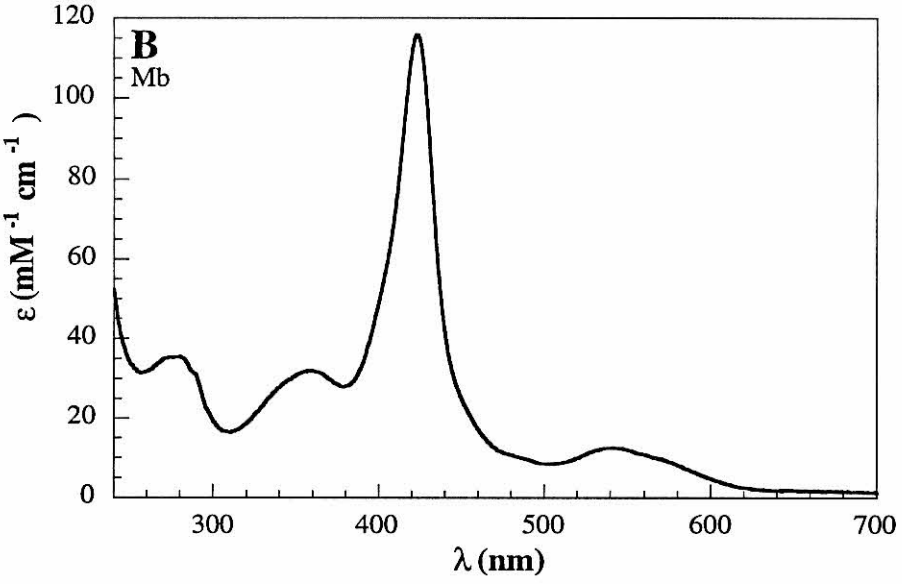

B
Mb
ε (mM-1 cm-1)
λ (nm)
0
20
40
60
80
100
120
300
400
500
600
700

Table 2.2

UV/vis absorption data for derivatives of horse heart Ala80cyt c and horse heart Mb

Protein	α	β	γ (Soret)	δ	UV
ferrous derivatives:					
deoxy-Ala80	548[a]		411/434[b]	*	277
	10.2		*87.2*[c]		*22.0*
deoxy-Mb	558		435	362[b]	280
	12.8		*121.0*	*37*	*34.4*
O_2-Ala80	571	538	406	348	278
	8.7	*11.0*	*113.2*	*25.0*	*23.5*
O_2-Mb	582	543	417	350	281
	15.4	*14.8*	136.5	*29.8*	*34.4*
CO-Ala80	528–560; broad		413	342	270
	~ 13		*233.6*	*27.3*	*31.3*
CO-Mb	577	542	423	346	273
	13.5	*15.7*	*183.3*	*30.8*	*41.7*
ferric derivatives:					
H_2O-Ala80 (pH 4.5)	622	495	400	353[b]	279
	3.7	8.1	*164.4*	*~ 31*	*22.2*
H_2O-Mb (pH 7.0)	634	504	409	360[b]	280
	5.0	*11.6*	*188*	*~ 39*	*36.5*
HO-Ala80 (pH 7.0)	622 569[b]	538	405	352	275
	1.6 *6.8*	*9.5*	*121.7*	28.1	*22.5*
HO-Mb (pH 10.0)	582 542	488[b]	412	352[b]	281
	10.1 *10.8*	*~11*	*125.2*	*~ 38*	*38.9*
CN-Ala80	542		414	357	273
	10.5		*99.5*	*28.9*	*19.4*
CN-Mb	541		423	360	280
	12.6		*115.9*	*32.0*	*35.6*

[a]Upright figures indicate λ(nm), italics indicate ε ($mM^{-1}cm^{-1}$).

[b]shoulders.

[c]ε at 411 nm.

* not resolved.

minor component that appears as a shoulder at 430 nm. Addition of exogenous ligands (CO, O_2) leads to the formation of products with single, sharp Soret bands; thus, the minor component is not believed to be a contaminant. Rather, the two components are proposed to be high-spin (λ_{max} ~ 430 nm) and low-spin (λ_{max} ~ 411 nm) heme.

The oxidized form of Ala80cyt *c* at pH 7 has spectroscopic properties significantly different from those of metMb. At pH 7, Fe(III)Mb has an absorption spectrum with an intense (ε = 188 $mM^{-1}cm^{-1}$) Soret band at 409 nm and weaker bands in the α, β region. The position, intensity, and shape of the h. h. Fe(III)Ala80 absorption bands are significantly different, suggesting different axial ligation. The absorption spectrum of metMb at pH 7 is characteristic of a high-spin heme with His-H_2O coordination.[2,37,38] However, the absorption spectrum of Fe(III)Ala80, in particular the position and intensity of the Soret band, is characteristic of a low-spin heme.[2,37,39] Increasing the pH of the metMb sample, however, results in deprotonation of the heme-bound water which triggers a transition from high-spin to low-spin; the solution turns from brown to red and the absorption spectrum becomes similar to that of Fe(III)Ala80cyt *c* at pH 7. Conversely, lowering the pH of the h. h. Fe(III)Ala80 sample below 6 results in a transition to a high-spin heme with a spectrum very similar to that of myoglobin at neutral pH (Figure 2.21). The pH-dependent equilibria of the highly analogous *Saccharomyces cerevisiae* Fe(III)Ala80cyt *c* mutant are discussed in depth in Chapter 4 of this thesis.

Other ligands were found to bind Ala80cyt *c* to give spectra similar to those seen for Mb. In the ferric state, CN^- binds strongly to Ala80cyt *c*, giving a species stable over months with a spectrum very similar to that of the CN-Mb species (Figure 2.22). Exposure of a reduced, deoxygenated sample of Ala80 to CO gives a derivative with spectral features like those of CO-Fe(II)Mb (Figure 2.19). The CO derivative was found to be very stable over time; the UV/vis spectrum of CO-Fe(II)Ala80cyt *c* was unchanged over many months.

In all of the Ala80 derivatives, the spectral features were observed to be blue-shifted by 5–10 nm from those of analogous Mb derivatives. This systematic difference, seen in both the oxidized and reduced forms of the proteins, is explained by the difference between the *c*-type heme in the cytochrome and the *b*-type heme in Mb. The Mb heme has vinyl groups where the cytochrome heme has saturated thioether substituents; this slightly different extent of conjugation accounts for the spectral shift.[30]

Electron Paramagnetic Resonance Spectroscopy

EPR is a valuable probe of the oxidation state and spin state of metal ions.[40] In the case of low-spin Fe(III) hemes, the three **g** values measured by EPR provide information on the ligand-field energies of the system. The measured **g** values for Fe(III)Ala80 and Fe(III)cyt *c* at pH 7 are presented in Table 2.3 along with published **g** values for selected low-spin hemes. The EPR spectra of h. h. Fe(III)Ala80 and Fe(III)cyt *c* are shown in Figure 2.23. The observation of three unequal **g** values for low-spin ferric hemes indicates that the three t_{2g} orbitals are of unequal but similar energy. This inequivalence is due to the asymmetry of the ligand field which is determined by the

Table 2.3

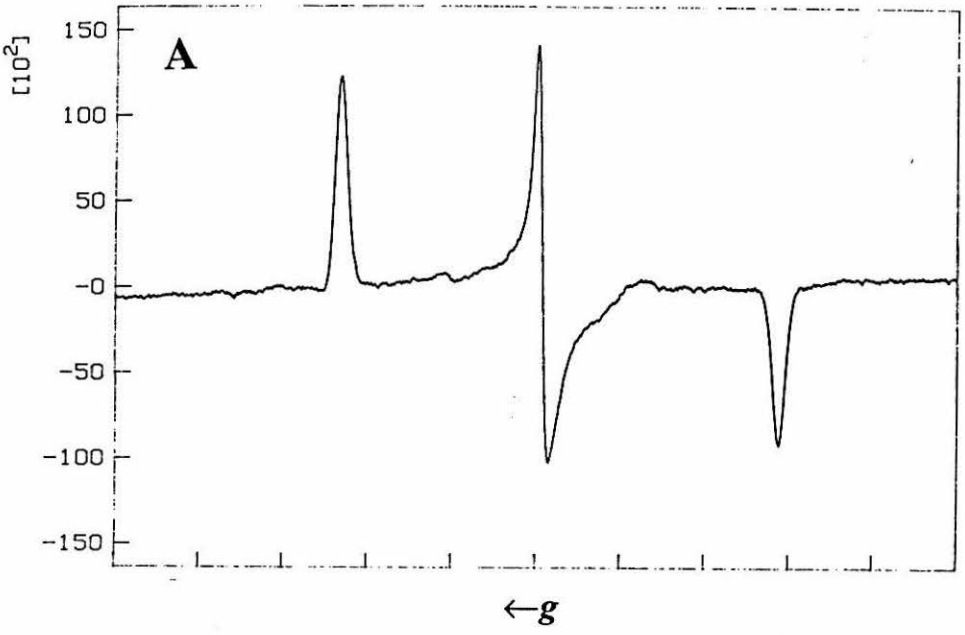
EPR parameters for horse heart Ala80cyt c and other low-spin ferric heme proteins

Protein	g_x	g_y	g_z	axial ligation	ref
Ala80cyt *c*	1.85	2.18	2.58	$His\text{-}OH^-$	*a*
human Mb, pH 10	1.84	2.16	2.59	$His\text{-}OH^-$	41
h. h. cyt *c*	1.21	2.21	3.03	His-Met	*a*
h. h. cyt c^b	1.24	2.24	3.06	His-Met	42
"alkaline" cyt *c*, pH 10	1.50	2.06	3.35	His-Lys	43
cytochrome b_5	1.41	2.22	3.05	His-His	44

[a]This work.

[b]Data acquired at 20 °C in 100 mM NaP_i (pH 7.6).

Figure 2.23 EPR spectra of (A) h. h. Fe(III)Ala80cyt *c* at 33 K and (B) native h. h. Fe(III)cyt *c* at 6.8 K. Samples were in 50 mM HEPES (pH 7.0), 50% glycerol.


A
150
100
50
-0
-50
-100
-150
[10²]
←g

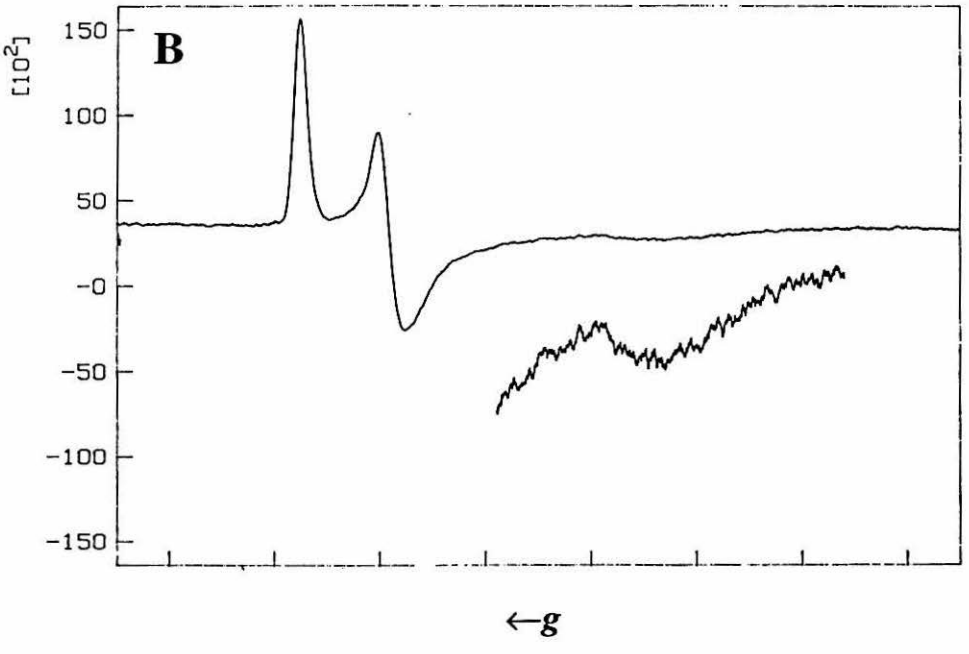

B
150
100
50
-0
-50
-100
-150
[10²]
←g

Figure 2.24 Energy levels of the five d orbitals of low-spin Fe(III), showing the splitting pattern of the orbitals as symmetry is reduced from cubic. The energy parameters V and Δ are indicated. Adapted from reference 40.

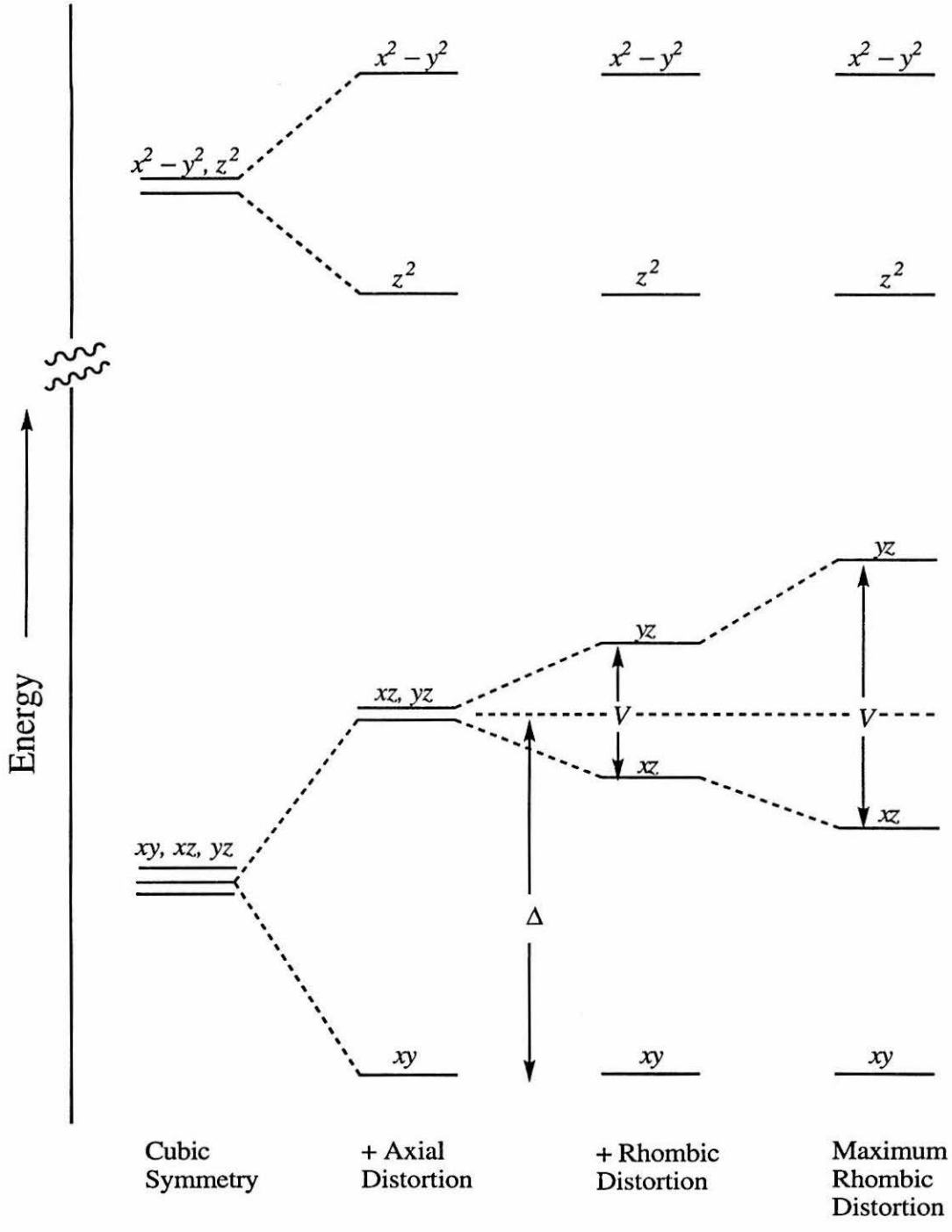
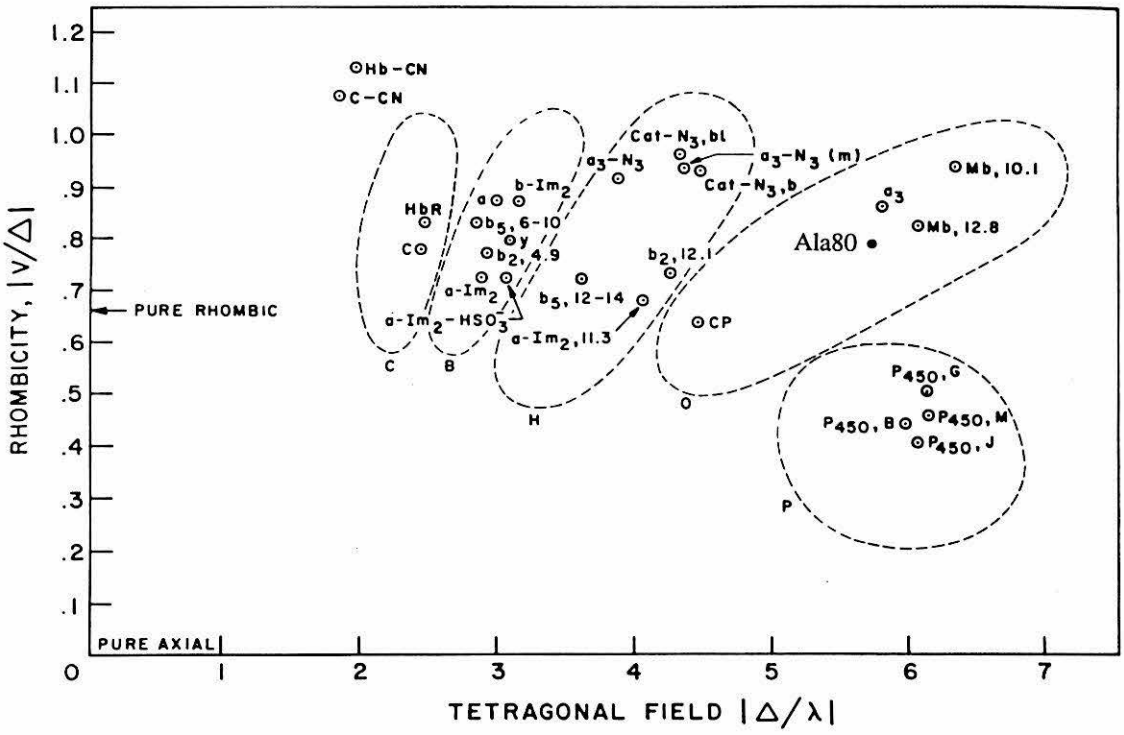

Energy
$x^2 - y^2$
$x^2 - y^2$
$x^2 - y^2$
$x^2 - y^2, z^2$
z^2
z^2
z^2
yz
yz
xz, yz
V
V
xz
xz
xy, xz, yz
Δ
xy
xy
xy
Cubic Symmetry
+ Axial Distortion
+ Rhombic Distortion
Maximum Rhombic Distortion

Figure 2.25 Correlation diagram for ligand-field parameters of low-spin Fe(III) heme proteins. Modified from reference 40 to show Fe(III)Ala80cyt *c*.

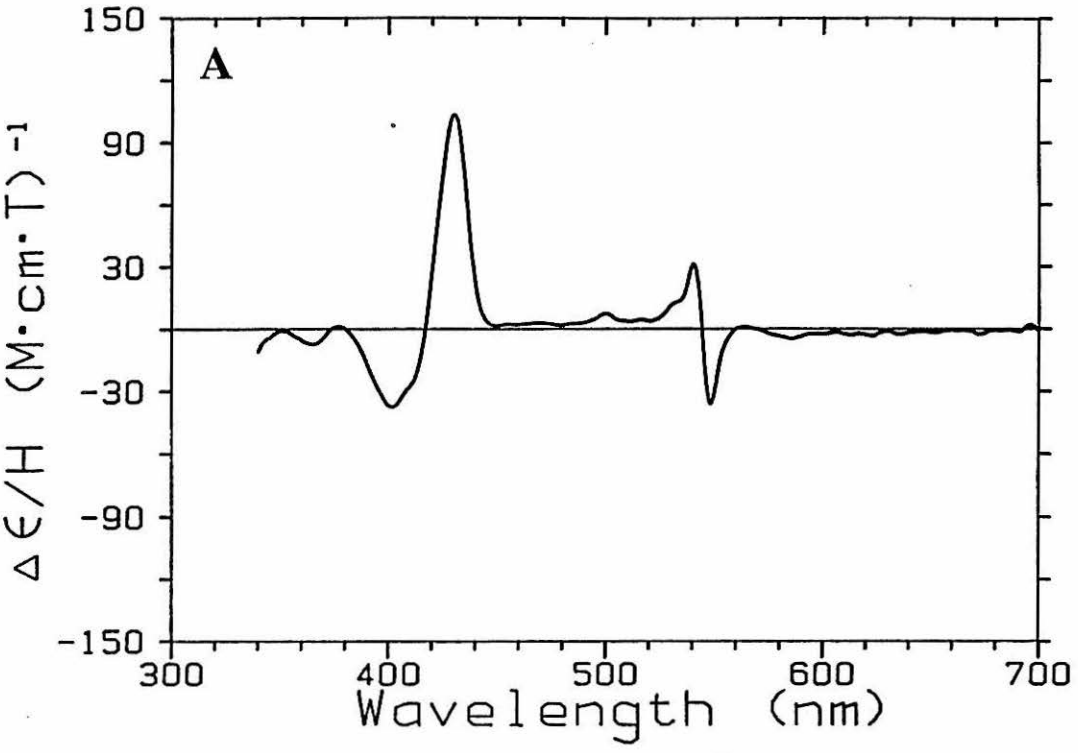

RHOMBICITY, |V/Δ|
TETRAGONAL FIELD |Δ/λ|
PURE RHOMBIC
PURE AXIAL
Hb-CN
C-CN
HbR
C
a
b-Im2
b5, 6-10
y
b2, 4.9
a-Im2
a-Im2-HSO3
b5, 12-14
a-Im2, 11.3
a3-N3
Cat-N3, bl
a3-N3 (m)
Cat-N3, b
b2, 12.1
CP
a3
Mb, 10.1
Mb, 12.8
Ala80
P450, G
P450, B
P450, M
P450, J
C
B
H
O
P

character of the two heme axial ligands.[32] Thus, the **g** values contain information on heme axial ligation. Equations 2.2–2.10 were used to relate the **g** values of h. h. Fe(III)Ala80cyt *c* to the relative values of the ligand-field energy parameters Δ and V. This analysis allows the description of the extent of distortion from cubic symmetry according to the energy parameters. The axial parameter, Δ, describes the extent of destabilization of d_{xz} and d_{yz} with respect to d_{xy}. The rhombic parameter, V, describes the stabilization of d_{xz} with respect to d_{yz} (Figure 2.24). This treatment only allows the calculation of these parameters as dimensionless ratios, Δ/λ and V/λ, where λ is the spin-orbit coupling constant. However, the relative values are useful for determining the nature of the heme axial ligands. The convention used is to plot the rhombicity, $|V/\Delta|$ vs. the tetragonal field, $|\Delta/\lambda|$. It has been observed that the results from a large body of data fall into five major domains.[33] On this correlation diagram, Ala80cyt *c* falls easily into the region corresponding to hemes with His-OH^- ligation (Figure 2.25).

Magnetic Circular Dichroism Spectroscopy

MCD is observed when an applied magnetic field perturbs the electronic states of a molecule, resulting in differential absorption of right- and left-circularly polarized light.[45] MCD has been shown to be particularly useful for probing heme oxidation and spin state as well as heme ligation[45cd] including axial ligand changes introduced by protein engineering.[34] The MCD spectrum of Fe(II)Ala80 exhibits features at 410 and 430 nm indicative of high-spin ferrous heme, as seen by the comparison to the spectrum of Fe(II)Mb (Figure 2.26A,C). In addition, the spectrum of deoxy-Fe(II)Ala80 has a small component in the α, β region (530–560 nm) that is a signature for a low-spin ferrous heme. An analogous band is observed in the MCD of Fe(II)cyt *c* (Figure 2.26).[45a] Comparison of the MCD spectra of Fe(III)Ala80 at pH 4.5 to published data for native and mutant metMbs confirms that Fe(III)Ala80 is high-spin at this pH (Figure 2.27).[45a] The presence of the intense feature near 400 nm in the spectrum of Fe(III)Ala80 at pH 4.5

Figure 2.26 (A) Measured MCD spectrum of h. h. deoxy-Fe(II)Ala80cyt *c* (8.9 μM protein in a 1-cm pathlength microcell) in 25 mM NaP_i (pH 7.0). (B) MCD spectrum of Fe(II)cyt *c*. Compare the low-energy band to that seen in the Ala80 spectrum; this is indicative of a small amount of low-spin protein for the mutant. (C) MCD spectrum of deoxy-Mb. The transition in the Soret region is similar to that seen for deoxy-Fe(II)Ala80 and indicates high-spin Fe(II). The spectra in (B) and (C) are from reference 45a.


A
150
90
30
-30
-90
-150
300
400
500
600
700
Wavelength (nm)
Δε/H (M·cm·T)⁻¹

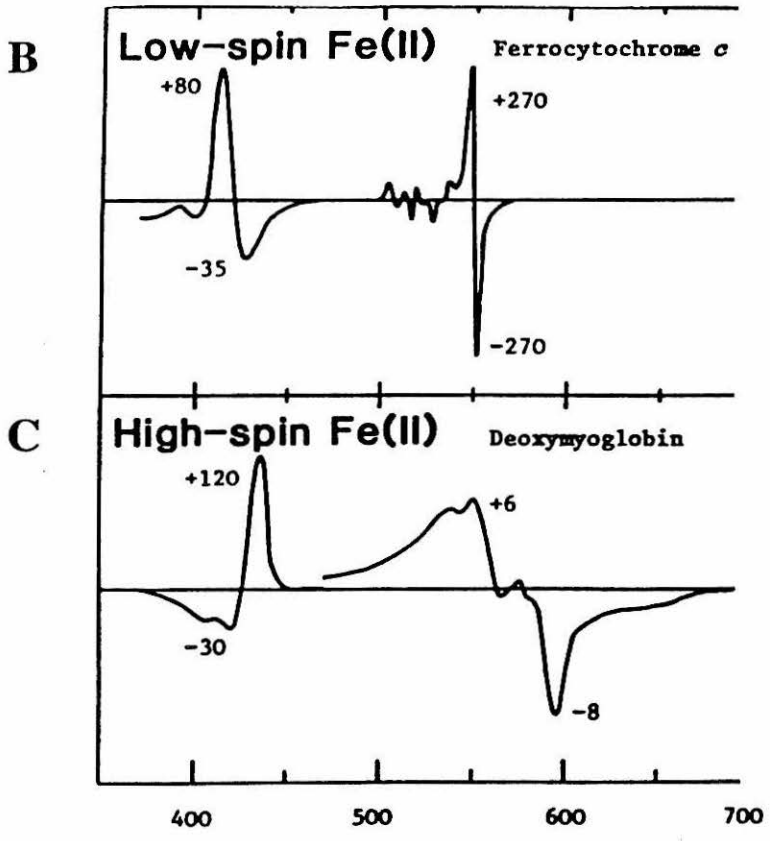
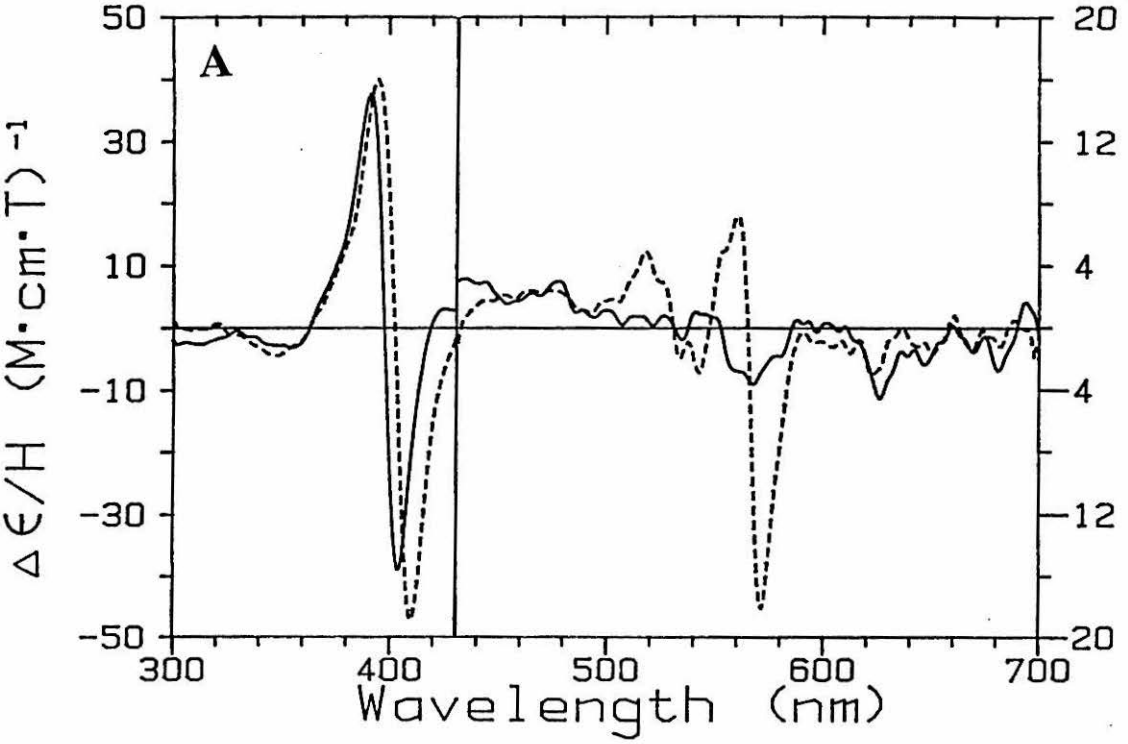
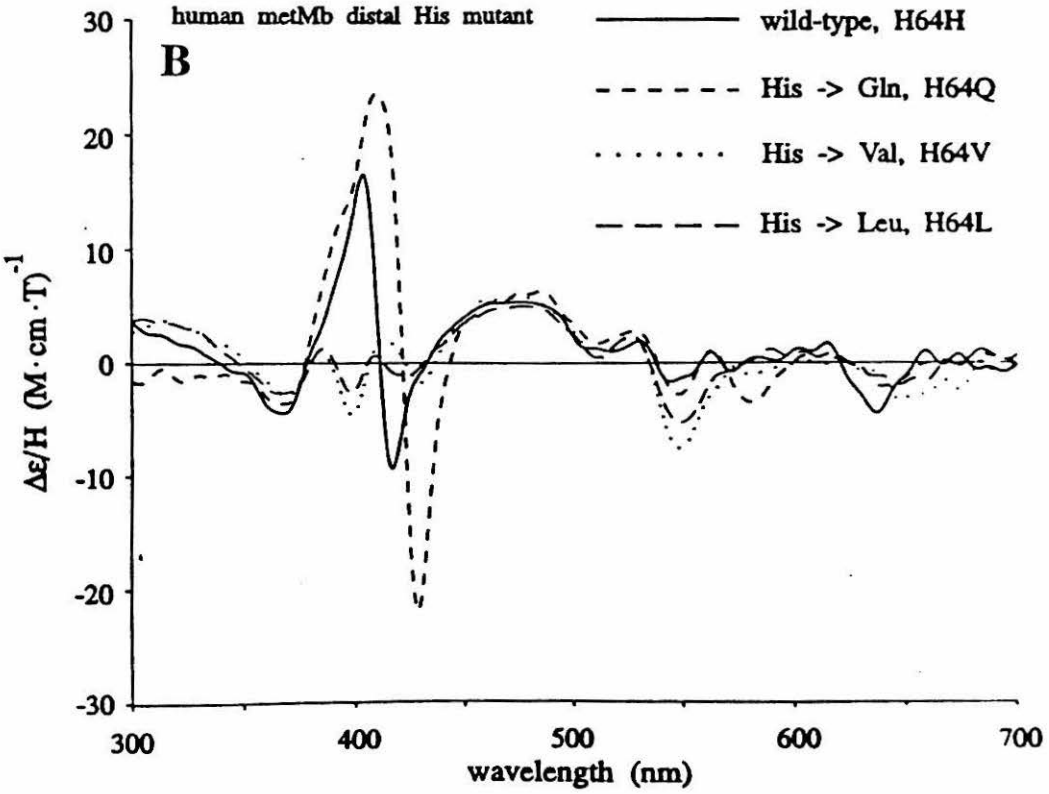

B
Low-spin Fe(II)
Ferrocytochrome c
+80
+270
-35
-270
C
High-spin Fe(II)
Deoxymyoglobin
+120
+6
-30
-8
400
500
600
700

Figure 2.27 (A) Measured MCD spectra of h. h. Fe(III)Ala80cyt *c* at pH 7 (dotted line) and pH 4 (solid line). The samples (45 μM in a 0.2-cm pathlength cell for pH 7, 42.8 μM in a 1-cm pathlength microcell for pH 4) were in 25 mM NaP_i (pH 7) or 25 mM NaOAc (pH 4). (B) MCD spectra of wild-type and distal His mutant human Fe(III)Mbs from reference 41. The His 64 → Val and His64 → Leu mutants do not have water coordinated, as indicated by the lack of a large transition in the Soret region.


A
50
30
10
-10
-30
-50
20
12
4
-4
-12
-20
300
400
500
600
700
Wavelength (nm)
Δ∈/H (M•cm•T)⁻¹


human metMb distal His mutant
B
wild-type, H64H
His -> Gln, H64Q
His -> Val, H64V
His -> Leu, H64L
30
20
10
0
-10
-20
-30
300
400
500
600
700
wavelength (nm)
Δε/H (M·cm·T)⁻¹

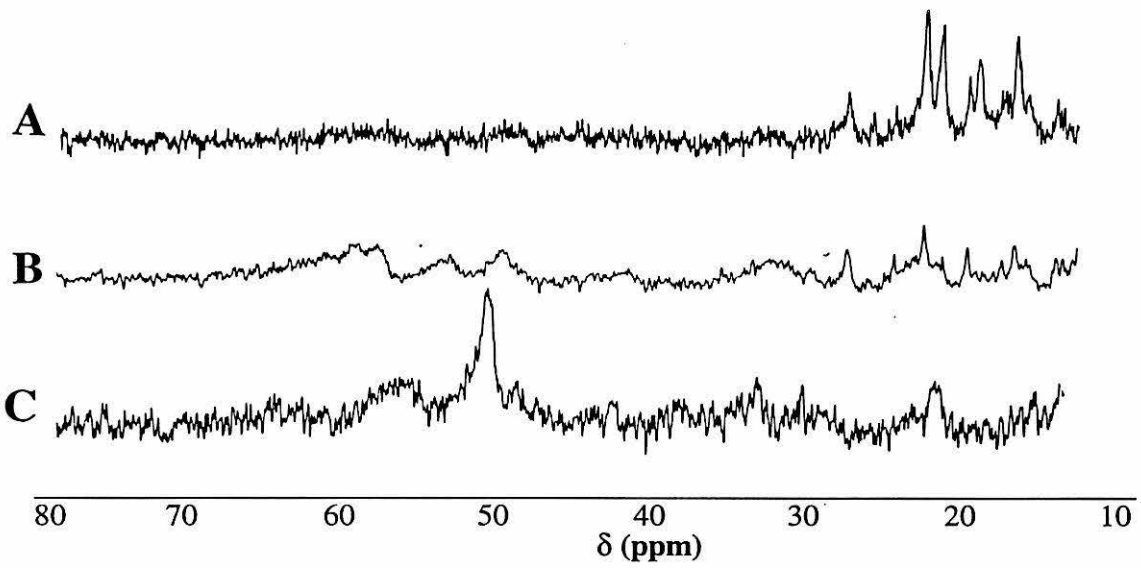
demonstrates that water does bind to the heme under these conditions to give a hexacoordinate iron; Mb mutants (His64Val and His64Leu) that are five-coordinate lack any intense MCD features in this region (Figure 2.27B).[41] The neutral pH form of semisynthetic Fe(III)Ala80 exhibits an intense feature between 550 and 580 nm typical of low-spin heme proteins, in agreement with the UV/vis and EPR data.

The MCD data thus confirm the hypothesis based on UV/vis spectroscopy that h. h. deoxy-Fe(II)Ala80 has a mixture of high- and low-spin heme. The physical basis for the mixture of spin states is not understood. It is possible that an axial ligand occupies the sixth coordination site of the heme part of the time to give a sample that is partly low-spin. Candidates for this ligand are water, Tyr67, and the nearby residue Lys79. It is unlikely that binding by water or Tyr, however, would cause the iron to be low-spin as both are weak-field ligands. A second possibility is that the constrained geometry of the heme site (Figure 2.1) forces the iron into the plane of the heme, causing the porphyrin to act as a stronger ligand. It is well-known that changes in spin-state in both ferrous and ferric hemes induce changes in the heme geometry and specifically the distance from the iron to the heme plane;[12] it is possible that a constrained geometry could affect some control on the iron spin state.

Paramagnetic Nuclear Magnetic Resonance Spectroscopy

The exquisite sensitivity that the shifts and linewidths of hyperfine-shifted resonances in the NMR spectra of paramagnetic heme proteins makes ^{1}H NMR a valuable technique for probing the electronic properties of Fe(III)Ala80cyt *c* in solution (see Chapters 4–6 for more details). The ^{1}H NMR spectrum of h. h. Fe(III)Ala80 recorded at pH 7.0 is exemplified by hyperfine-shifted resonances in the 10–30 ppm region (Figure 2.28A), characteristic of low-spin Fe(III).[46] This confirms that the neutral pH form of Fe(III)Ala80 in solution at room temperature is low-spin, as inferred from absorption spectroscopy. As the pH is decreased to 5.2, a number of broad signals appear

Figure 2.28 Downfield region of the 500 MHz ^{1}H NMR spectrum of h. h. Fe(III)Ala80cyt *c* at (A) pH 7.0, (B) pH 5.2, and (C) pH 3.8. The sample was 0.25 mM in protein in 25 mM NaP_i in D_2O. The residual HDO peak was suppressed by presaturation. The reported pH values are not corrected for the isotope effect.


A
B
C
80
70
60
50
40
30
20
10
δ (ppm)

in the 30–70 ppm region, characteristic of high-spin Fe(III) (Figure 2.28B).[46] A high-spin Fe(III) species becomes dominant as pH is lowered; at pH 3.8 the spectrum is characterized by very broad signals indicative of high-spin Fe(III) (Figure 2.28C).

Upon comparison of the h. h. Fe(III)Ala80 spectrum with that of native h. h. cyt *c* (Figure 2.29), it is immediately apparent that the pattern and linewidth of the hyperfine-shifted signals is drastically altered by the Met80Ala mutation. The larger linewidth seen for the hyperfine-shifted Ala80 signals is consistent with its heme being ligated by HO^-, which is a weaker ligand than the Met thioether. The spectrum of CN-Fe(III)Ala80 (Figure 2.30) in fact has very narrow hyperfine-shifted signals, consistent with ligation by a strong-field ligand. The change in the pattern of hyperfine shifts upon mutation of Met80 to Ala is also consistent with the removal of Met80 from heme coordination. Most of the hyperfine-shifted resonances in spectra of heme proteins are due to heme substituents, such as the heme methyls. Because the contact contribution to the hyperfine shift of heme methyls is significant, the pattern of heme methyl shifts depends mainly on the unpaired spin density at the methyls.[46,47] The asymmetry of spin delocalization on the heme for His-Met-ligated cytochromes, in turn, is determined by the orientations of the His and Met ligands.[48,49] Removal of the Met from coordination thus drastically changes the pattern of hyperfine-shifted resonances, as is observed here in both the spectra of HO-Ala80cyt *c* and CN-Ala80cyt *c*.

Ligand Photolysis

The rates of ligand binding to heme proteins subsequent to photodissociation provide information about the reactivity of the heme and the structure of the heme pocket.[12,23-27] The rate constants for recombination of O_2 and CO to h. h. Fe(II)Ala80 and to h. h. Fe(II)Mb are summarized in Table 2.4. Measured rates for geminate and bimolecular recombination of CO and O_2 to Mb agree well with literature values.[50] The

Figure 2.29 Downfield region of the 500 MHz ^{1}H NMR spectrum of native h. h. Fe(III)cyt *c*. The sample was 2 mM in protein in 25 mM NaP_i (uncorrected pH 7.0) in D_2O.

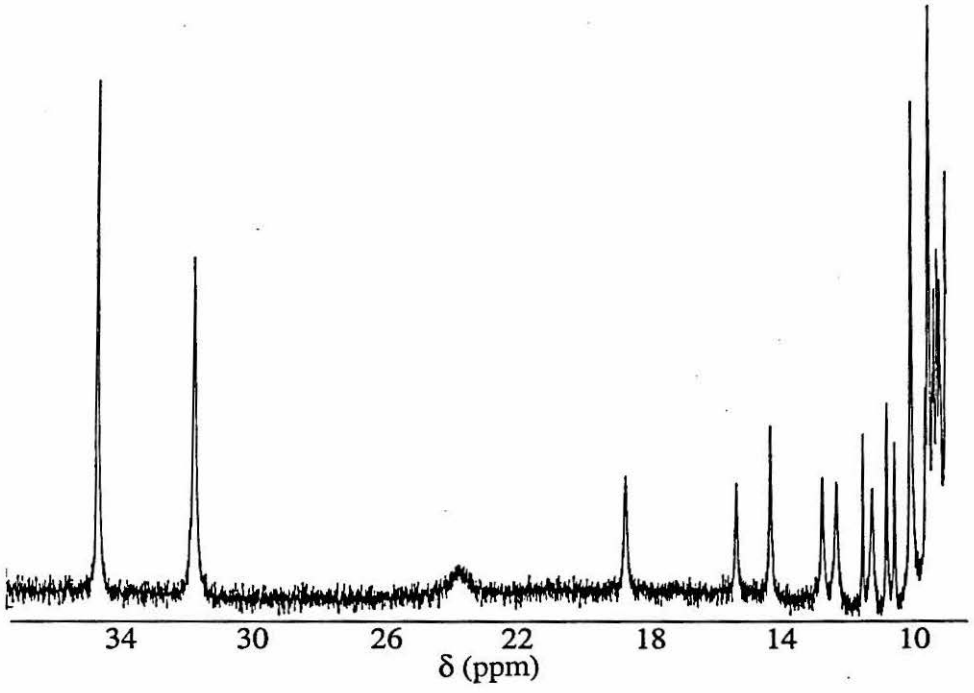
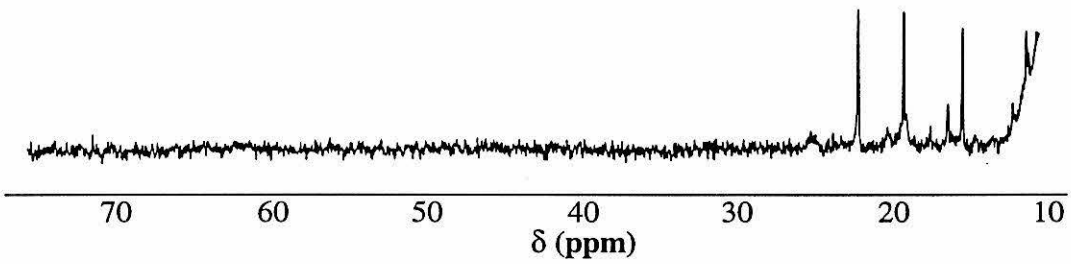

34
30
26
22
18
14
10
δ (ppm)

Figure 2.30 Downfield region of the 500 MHz ^{1}H NMR spectrum of the cyanide adduct of h. h. Fe(III)Ala80cyt *c*. The sample was 0.25 mM in protein in 25 mM NaP_i (uncorrected pH 7.0) in D_2O.


70
60
50
40
30
20
10
δ (ppm)

kinetics of ligand recombination with h. h. Ala80, however, are drastically different from those reported for Mb (Table 2.4).

The observed rate of CO binding to Fe(II)Ala80cyt *c* on timescales from 500–1000 ms was found to increase linearly with [CO], confirming that bimolecular recombination kinetics were observed. In this case, CO is photolyzed from the heme and diffuses through the protein into solution before rebinding. Data collected under 1 atm of CO were found to be most reliable for determination of k_{on} as they occur on a timescale amenable to measurement on the laser system used. In addition, the error in measurement of the amount of CO introduced over the sample is minimized; thus, these experiments were used to calculate the best value for k_{on}. The transient absorption data for CO binding to Fe(II)Ala80 and to Fe(II)Mb after photodissociation are presented in

Table 2.4

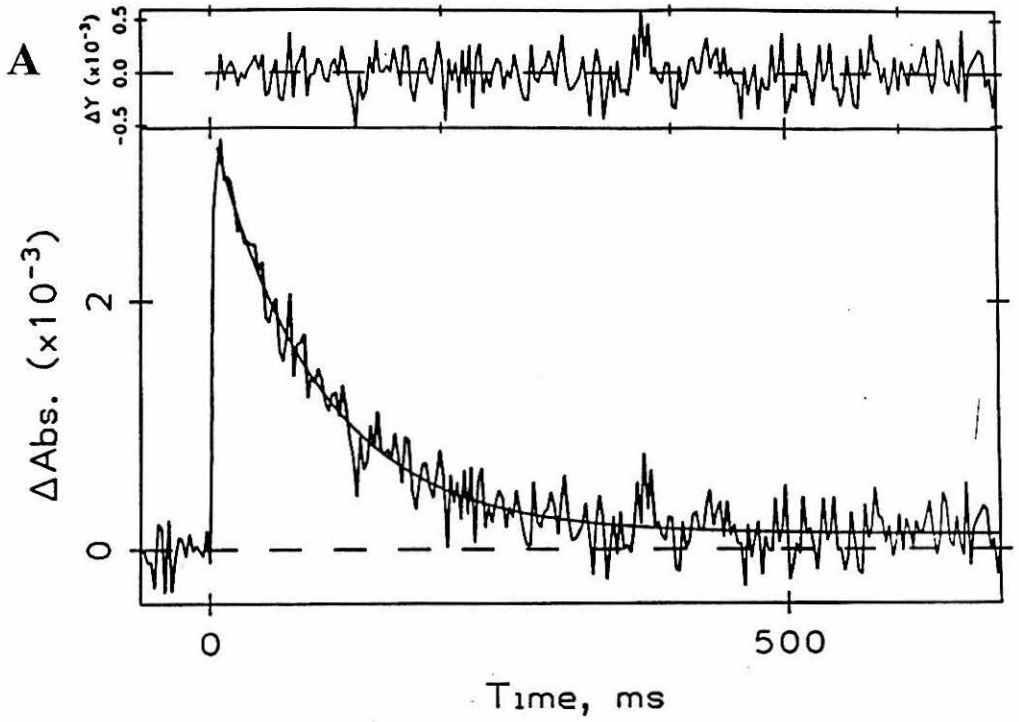
Rate constants for ligand binding to native and mutant myoglobins and to horse heart Ala80cyt c

Protein	k^{gem}_{O2} x 10^{-6} s^{-1}	k^{bi}_{O2} x 10^{-6} $M^{-1}s^{-1}$	k^{bi}_{CO} x 10^{-6} $M^{-1}s^{-1}$	ref
Horse heart Mb	7.3	5.8	0.30	*a*
Sperm Whale Mb	9.6	17	*	53
Sperm Whale Mb	*	16	0.55	54
L29F	*	21	0.22	54
H64G	*	140	5.8	20a
V68F	*	1.2	.25	55
Aplysia Mb	*	24	0.65	56
Ala80cyt *c*	*	app. 0.002	0.0087	*a*

[a]this report.

*not determined or reported.

Figure 2.31 Transient absorption kinetics of the recombination of CO to (A) h. h. Fe(II)Ala80cyt *c* and (B) h. h. Fe(II)Mb after photolysis. The decays correspond to the decrease in absorption of the Soret band of the deoxy species as CO binds. The smooth lines indicate fits to a single exponential function under pseudo-first-order conditions. For these experiments, [CO] = 1 mM and (A) $k_{ps} = 8.0\ s^{-1}$; (B) $k_{ps} = 300\ s^{-1}$.


A
ΔY (x10⁻³)
0.5
0.0
-0.5
ΔAbs. (x10⁻³)
2
0
0
500
Time, ms

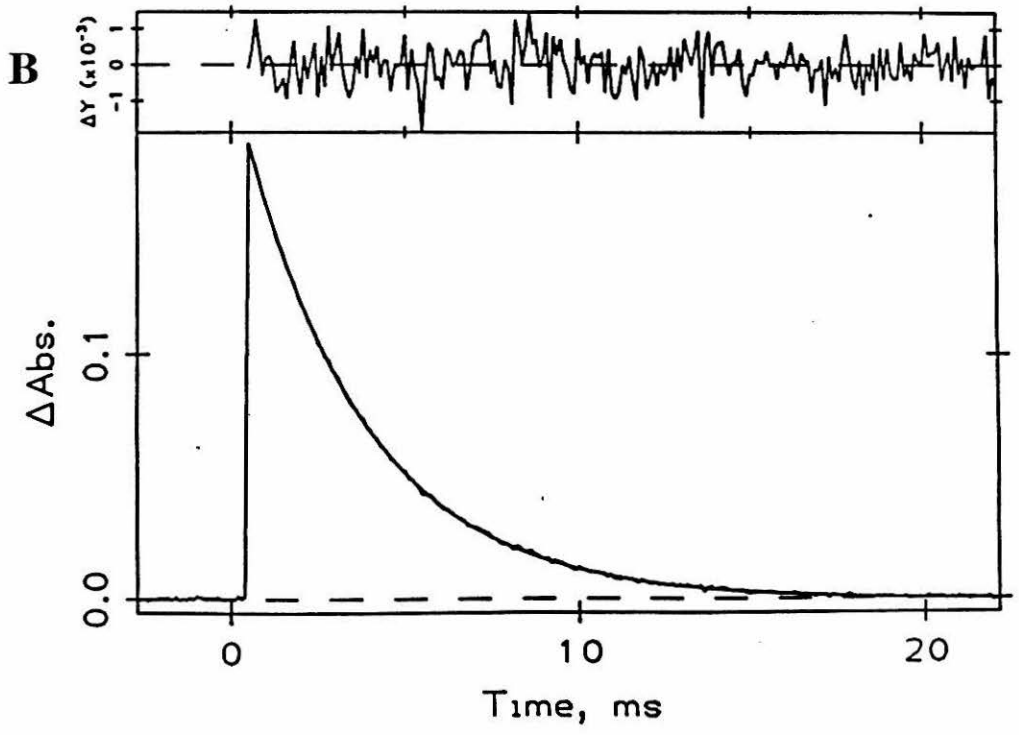

B
ΔY (x10⁻³)
1
0
-1
ΔAbs.
0.1
0.0
0
10
20
Time, ms

Figure 2.31. An analysis of eight data sets gave second-order rate constants k_{on} = 8.7(8) x 10^3 $M^{-1}s^{-1}$ for h. h. Ala80cyt *c* and k_{on} = 3.0 x 10^5 $M^{-1}s^{-1}$ for h. h. Mb. The quantum yield of photodissociation of CO from h. h. Fe(II)Ala80 relative to Mb, Φ_{rel}, calculated as described in Materials and Methods of this chapter, was found to be 0.02. It has been proposed that for all CO derivatives of ferrous heme proteins, the quantum yield of photodissociation is unity on the femtosecond timescale and that the observed quantum yield depends on the behavior of the protein on the timescale of observation.[23] In the case of Ala80cyt *c*, this implies that most CO rebinds geminately from within the protein matrix on very short timescales. No evidence of geminate recombination to h. h. Fe(II)Ala80 was observed using the nanosecond transient absorption system.

Photodissociation studies were also attempted on O_2-Fe(II)Ala80cyt *c*. Very small signals were observed; Φ_{rel} for O_2 photodissociation was found to be $\Phi_{rel} < 0.01$, and the resulting small signal size precluded kinetics studies. In the case of Mb, O_2 photodissociation typically proceeds with yields of 0.1–0.5.[51]

The low yields of photoejection of ligands from the Ala80cyt *c* polypeptide matrix, along with the slow rate observed for CO recombination, indicate that the h. h. Fe(II)Ala80 heme is much less accessible to small molecules than the heme of Fe(II)Mb. This is in spite of the fact that the heme in cyt *c* is not more buried; in fact, the cyt *c* heme edge is slightly more solvent-exposed than that of Mb.[3a,13] It is not logical to attribute the difference in ligand-binding kinetics to Ala80 and Mb to the properties of the heme itself, in particular because changes in heme geometry that accompany ligand-binding typically occur on timescales of picoseconds or less,[26] while bimolecular ligand-binding occurs on timescales of milliseconds. The best explanation for this difference seen in CO binding kinetics and photodissociation yields is that the cytochrome fold is built in such a way as to not allow facile movement of ligands through the polypeptide matrix, while that of Mb is. The mechanism of ligand-binding to Mb remains unclear, but is believed to involve movement of amino acid side chains to make a ligand-binding channel and/or

a global "breathing" motion of the globin.[52] Apparently, these mechanisms are not available to the cytochrome fold. Ala80cyt *c*, therefore, would be a poor physiological substitute for Mb because the slow movement of ligands through the polypeptide would lead to very inefficient uptake and delivery of dioxygen in the body. As an electron transfer protein, it is advantageous to cyt *c* to prevent ligand-binding to the heme as this would interfere with facile and reversible electron transfer.

Autoxidation

The h. h. O_2-Fe(II)Ala80 visible absorption spectrum as it changes with time is presented in Figure 2.32A. A plot of $(A_\infty - A_t)$ vs. time gives a curve that fits to a single-exponential function and indicates an autoxidation rate of 0.010(3) h^{-1} (Figure 2.32B). This very slow rate of autoxidation is significantly slower than that of most Mbs (Table 2.5). Because autoxidation of heme proteins requires the release of superoxide and the uptake of water (or hydroxide) for those proteins that have a sixth ligand in the ferric state, it is proposed that the stability of O_2-Fe(II)Ala80cyt *c* to autoxidation is related to the protection of the heme from the solvent. Hydrogen-bonding to the O_2-Fe(II) unit would also retard autoxidation.

Table 2.5

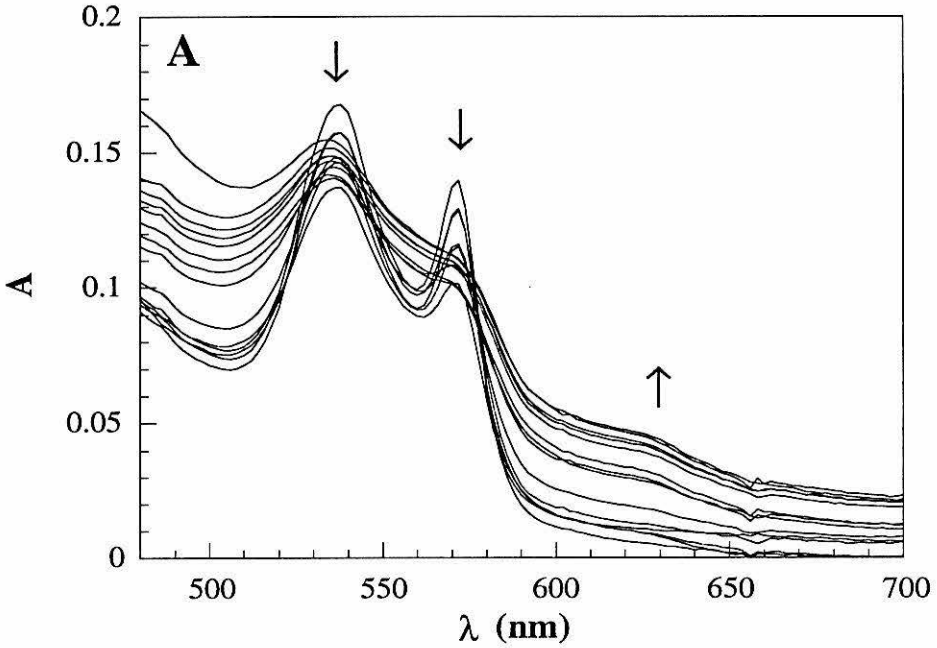
Autoxidation rates for horse heart O_2-Fe(II)Ala80cyt c and selected Mbs[a]

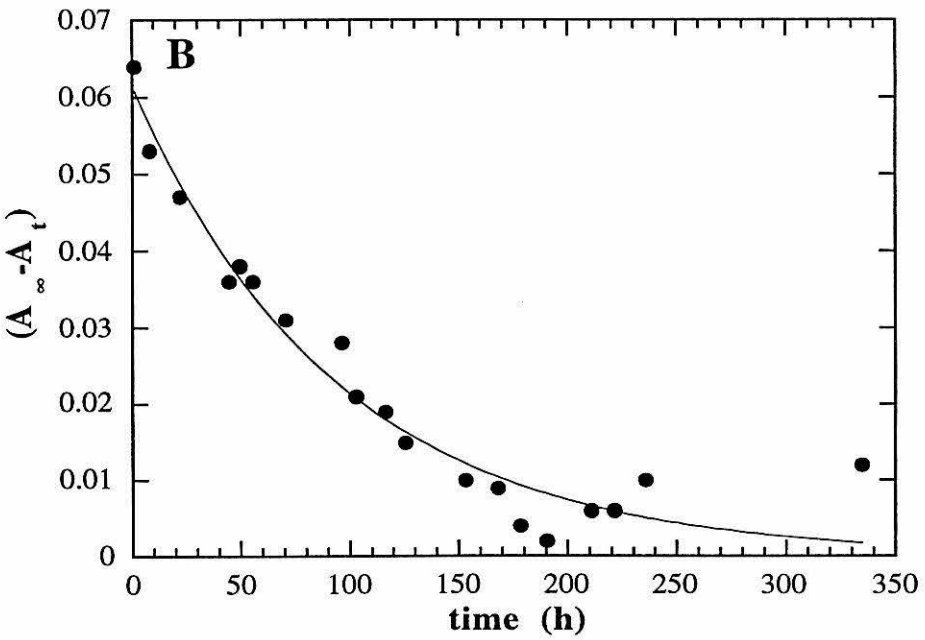
Protein	k_{ox}, h^{-1}	ref
Horse heart Ala80cyt *c*, pH 7.0	0.01	[b]
Yellowfin tuna Mb, pH 5.9	0.18	57
Sperm whale Mb, pH 5.78	0.08	57
Bovine Mb, pH 5.77	0.13	57

[a]Rates measured at 22 °C.

[b]This work.

Figure 2.32 (A) UV/vis absorption spectra taken over the course of the autoxidation of h. h. O_2-Fe(II)Ala80cyt *c*. Data were recorded on 15 μM protein in 25 mM NaP_i (pH 7.0), at 22 °C. The loss of isosbestic points indicates that other processes are occurring, presumably protein degradation. (B) Plot of autoxidation data. The quantity $(A_\infty - A_t)$ is plotted against time and the data are fit to a single exponential to obtain the autoxidation rate, 0.010(3) h^{-1}.


A
0.2
0.15
0.1
0.05
0
A
500
550
600
650
700
λ (nm)


B
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
(A ∞ -A t)
0
50
100
150
200
250
300
350
time (h)

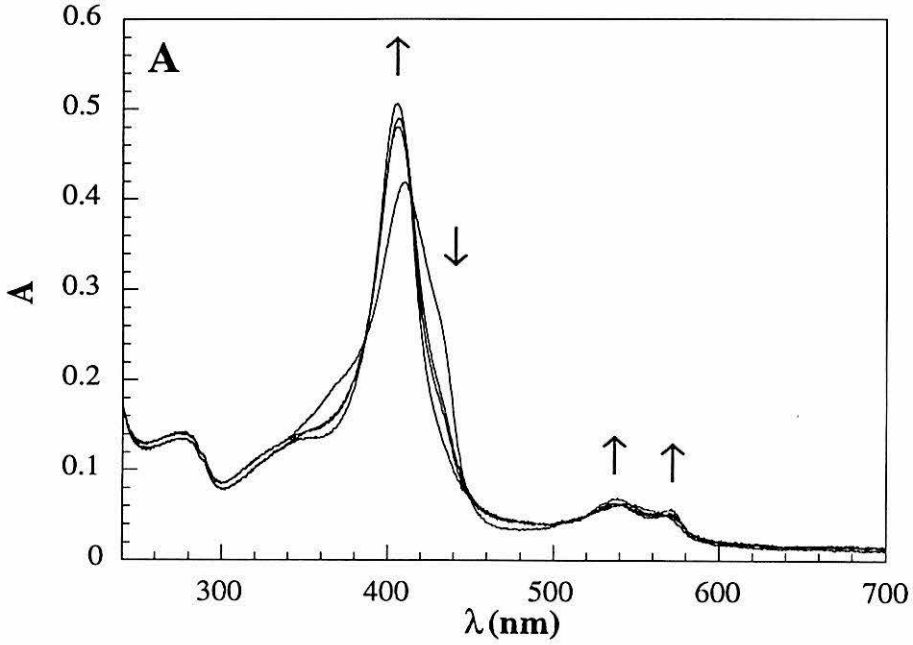
As is apparent in Figure 2.32A, isosbestic points are not maintained throughout this autoxidation reaction. In contrast, the autoxidation of O_2-Fe(II)Phe67Ala80 proceeds with clean isosbestic points (*vide infra*). This suggests that the Tyr67 may play a role in the autoxidation reaction, forming radicals which subsequently perform deleterious reactions on the protein.

Dioxygen Affinity

The changes in the spectrum of Fe(II)Ala80cyt *c* as dioxygen is added (Figure 2.33A) are consistent with a two-state transition from deoxy-Fe(II)Ala80 to O_2-Fe(II)Ala80. Fitting the Hill equation (2.29) to the data as described in Materials and Methods for the four experiments performed gives a dioxygen association constant of $K =$ (2.6 ± 1.4) μM^{-1} (Figure 2.33B). The error was determined by applying standard error analysis to the data sets. Because of the extremely slow binding rate for ligands to h. h. Fe(II)Ala80, an equilibration time of one hour was needed for each point. In spite of the resistance of O_2-Fe(II)Ala80cyt *c* to autoxidation, at low P_{O_2} values the autoxidation rate is increased and protein autoxidation occurs during the course of the experiment if many points are collected. This unfortunately limits dioxygen affinity experiments for this system and the reported number should be considered an approximation. Because of this, studies comparing binding constants with other ligands and other mutants were not attempted because the results could be misleading.

In spite of the large error, it is clear that Fe(II)Ala80 has a high affinity for dioxygen, comparable with that of natural O_2-storage systems (Table 2.5). This is an extremely surprising result; it was not expected that a protein that evolved to perform electron transfer could bind dioxygen more strongly than do natural Mbs. The affinity of various species and variants of Mb for dioxygen has been linked to interactions between residues in the distal heme pocket with the bound dioxygen molecule.[12,15,16,58] In most species of Mb, the distal His is responsible for the stabilization of the ferrous-oxy species

Figure 2.33 (A) UV/vis spectra of h. h. Fe(II)Ala80cyt *c* taken at various concentrations of O_2 used for determining O_2 binding constant. The arrows indicate the spectral changes as P_{O_2} is increased. (B) Hill curve fit to data in (A).


A
0.6
0.5
0.4
0.3
0.2
0.1
0
A
300
400
500
600
700
λ(nm)

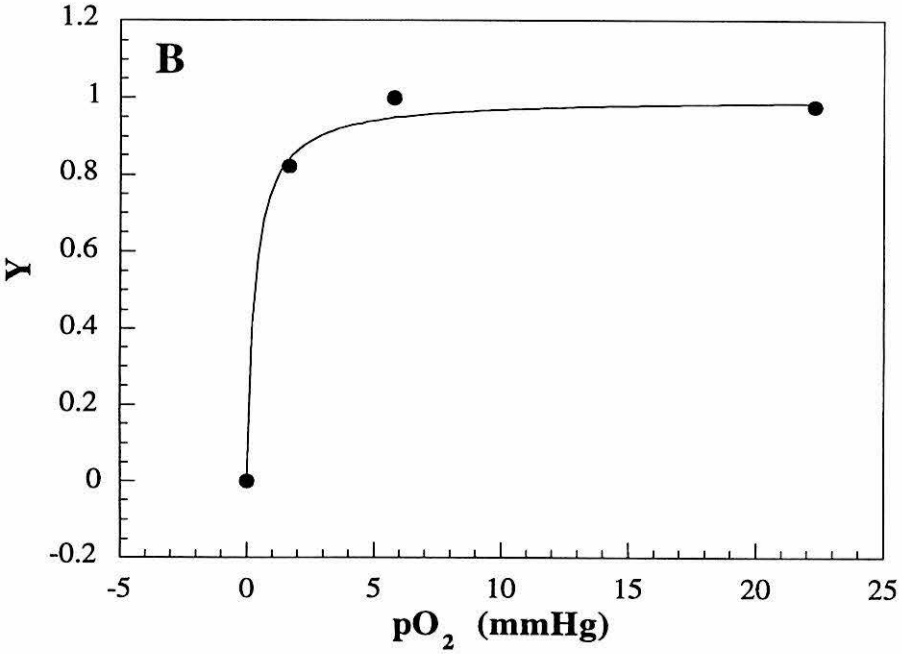

B
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
Y
-5
0
5
10
15
20
25
pO2 (mmHg)

Table 2.6

Dioxygen affinity of horse heart Ala80cyt c and selected myoglobins and hemoglobins

Protein	K, μM^{-1}	ref
Horse heart Ala80cyt *c*, pH 7.0	2.6	*a*
Horse heart Mb, pH 7.0	0.85	12
Human Mb I, pH 8.0	0.85	12
Sperm whale Mb, pH 8.6	1.13	12
Aplysia Mb, pH 7.0	0.20	12
Glycera Mb	0.10	15
Ascaris Hb, pH 6.5	76.3	58c

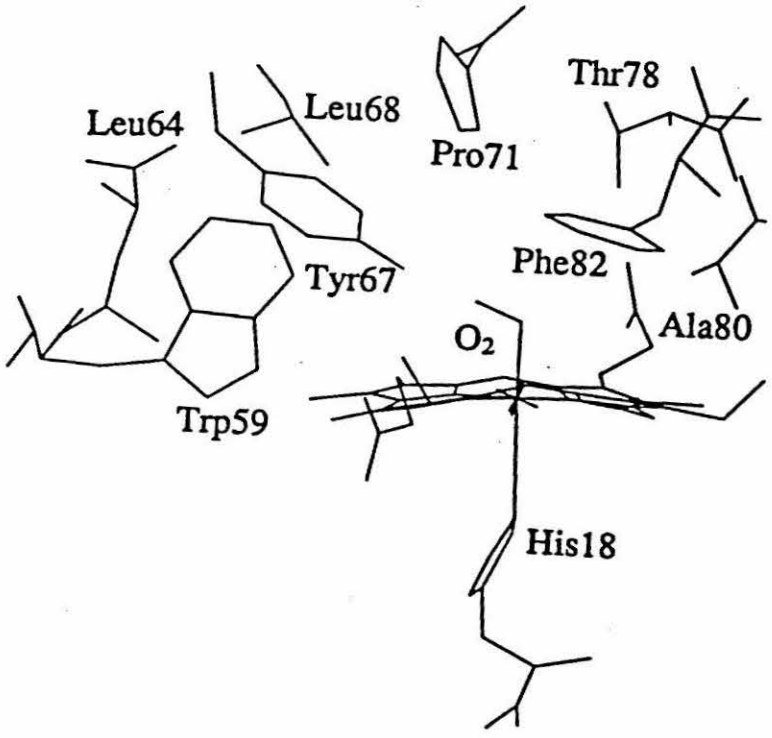
[a]This work.

through hydrogen-bonding. Other residues have also been implicated in dioxygen stabilization in various species of Mb and Hb[58], such as Arg (*Aplysia* Mb),[58a] Gln and Phe (Asian Elephant Mb),[58b] and Tyr (*Ascaris* Hb).[58c] Tyr67 in the distal pocket of Ala80cyt *c* is a logical candidate for a distal-His analog in Ala80cyt *c*.

Molecular Modeling of Horse Heart O_2-Fe(II)Ala80cyt *c* Active Site

To investigate what interactions may occur between the O_2 bound to Fe(II)Ala80 and residues in the distal pocket, molecular modeling of the h. h. O_2-Ala80cyt *c* active site was performed using the program Biograf. The Met80 side chain in the X-ray crystal structure of native h. h. cyt *c* was converted to a methyl group and all residues within a 10 Å sphere of the Ala80 side chain were energy-minimized using the molecular mechanics option. The heme was kept fixed in the minimization. A dioxygen molecule was then built and bound to the heme iron and it was minimized with respect to the heme; a Fe-O-O angle of 138° resulted, which compares well with that of 143° found in sperm whale Mb.[13] The Phe82 side chain was rotated from its position in the native h. h. cyt *c* crystal structure to interact with the dioxygen, and the Tyr67 was moved into position to

Figure 2.34 Molecular model of the h. h. O_2-Fe(II)Ala80cyt *c* distal pocket. Modeling was performed by using Biograf version 2.2 on a VaxStation 3500. The model was built on the 1.94-Å crystal structure of h. h. cyt c^{3a} by substitution of Ala for Met80, placement of dioxygen at the heme iron, and local energy minimization. Phe82 was rotated from its position in the native structure before local energy minimization to show the potential Phe-O_2 interaction.


Leu64
Leu68
Pro71
Thr78
Tyr67
Phe82
Ala80
O2
Trp59
His18

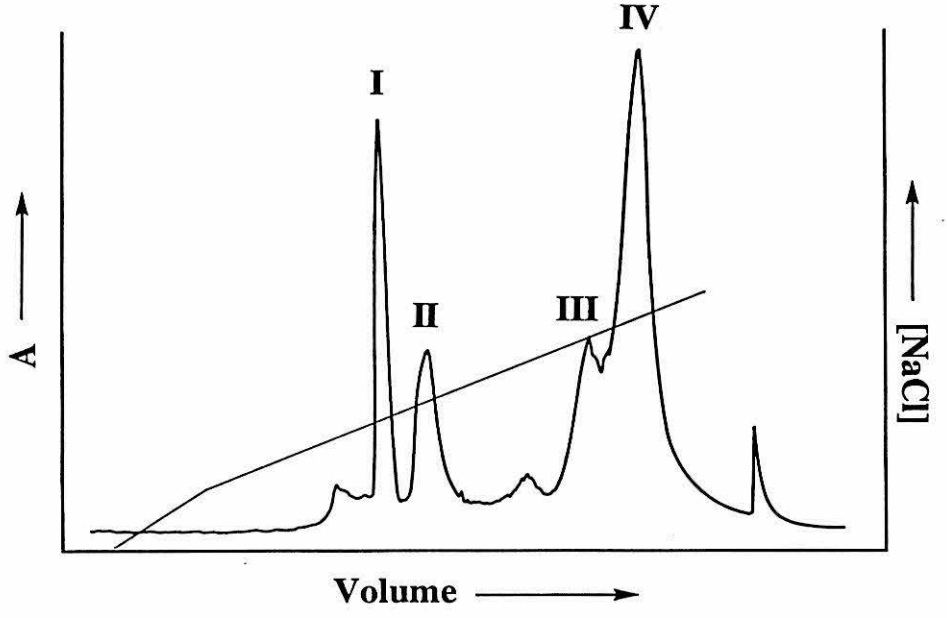
hydrogen-bond with the second O atom of the dioxygen molecule. Energy minimization was then performed one last time to correct for any unfavorable steric interactions. The resulting model is presented in Figure 2.34. In this model, the Phe82 Cζ → O2(2) distance is 3.4 Å, compared with 3.2 Å reported for Leu29PheMb,[20a] suggesting that the positive edge of Phe82 may interact favorably with bound dioxygen. It was also determined with this model that Tyr67 may hydrogen-bond with bound dioxygen; the Tyr67 Oη → O_2(2) distance is 2.65 Å in the model, compared with 2.77 Å from the sperm-whale Mb distal His Nε2 → O_2(2).[13]

Characterization of Horse Heart Phe67Ala80cyt *c*

Chromatography and Electrophoresis

The Phe67Ala80cyt *c* semisynthesis reaction mixture elutes on a cation exchange column with a profile (Figure 2.35) very similar to that of the Ala80cyt *c* reaction mixture. As in the case of Ala80cyt *c*, Fe(II)Phe67Ala80cyt *c* elutes first at 21 mM NaCl (peak I), followed by Fe(III)Phe67Ala80 at 23 mM NaCl (peak II); both are well-separated from unreacted cyt *c*(1–65). The reduced protein has a pink/orange color much like that of Fe(II)Ala80cyt *c*, but the oxidized protein is red-brown, suggesting that it has some high-spin character. The chromatographic and spectroscopic properties of the oxidized and reduced forms of semisynthetic Phe67Ala80cyt *c* are readily interconverted by chemical oxidation and reduction. Yields for Phe67Ala80 (~ 20%) are consistently higher than those seen for Ala80 reconstitution. It is possible that replacement of Tyr67 with the more hydrophobic Phe enhances hydrophobic interactions between the heme and the peptide in the reconstitution reaction. An alternative explanation for the higher reconstitution yield is that mutation of Tyr67 to Phe removes a potential heme ligand, the phenol side chain of Tyr, thus eliminating the production of a non-productive complex between cyt *c*(1–65) and (66–104) with a Tyr-ligated heme. Precedent for this has been observed in the failure to reconstitute His79Ala80cyt *c* and His72Ala80cyt *c*. For these

Figure 2.35 Typical cation-exchange (FPLC Mono S HR 10/10) chromatogram for the purification of h. h. Phe67Ala80cyt *c* from the reconstitution reaction. Peaks I and II are reduced and oxidized Phe67Ala80cyt *c*, respectively. Unreacted cyt *c*(1−65) elutes at higher ionic strength (peaks III and IV). The thin line represents the gradient of NaCl. Detection is at 280 nm on the 1.0 absorbance scale.


IV
I
II
III
A
[NaCl]
Volume

reactions, no cyt *c*-like product was formed and it is believed that coordination of the introduced His side chains to the heme prevents productive complex formation.[59] No residues besides Met80 known to bind strongly to hemes are found in the native sequence of cyt *c*(66−104) (Figure 2.4). SDS/PAGE and CZE analysis confirmed that a homogeneous product with a molecular weight and CZE mobility similar to native cyt *c* was formed.

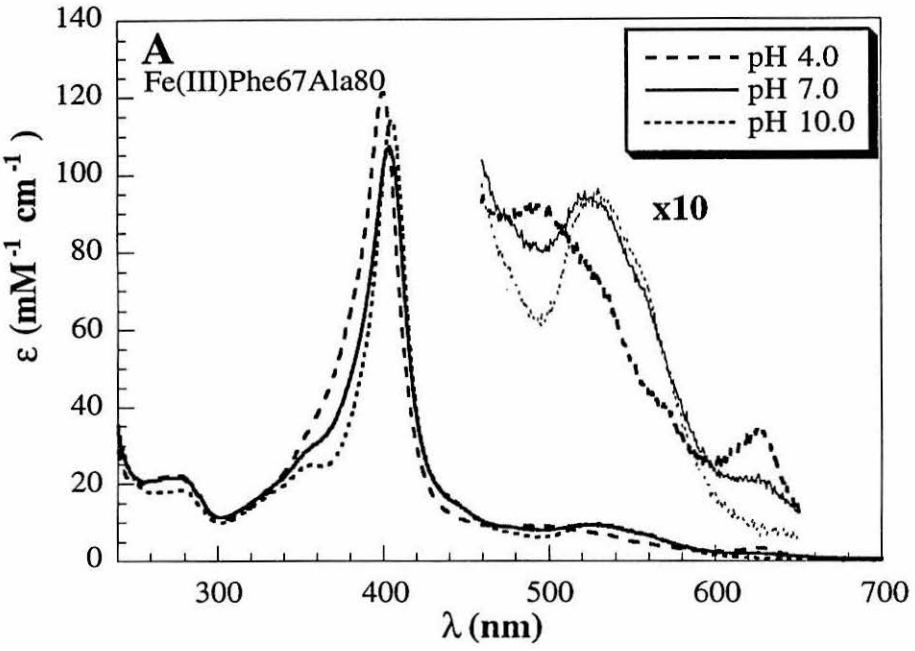
Absorption Spectroscopy

The UV/vis absorption spectra of oxidized (Figure 2.36A) and reduced (Figure 2.36B) h. h. Phe67Ala80 are in analogy to those of h. h. Ala80. The spectrum of Fe(II)Phe67Ala80 is very similar to that of h. h. O_2-Fe(II)Ala80, except that the α-band is slightly blue-shifted (568 nm vs. 571 nm) and the Soret band has a smaller extinction coefficient (97.8 $mM^{-1}cm^{-1}$ vs. 112.6 $mM^{-1}cm^{-1}$). The UV/vis absorption spectrum of O_2-Fe(II)Phe67Ala80, by its similarity to that of O_2-Fe(II)Ala80 and to O_2-Mb, suggests that dioxygen also binds Fe(II)Phe67Ala80. In the ferric state, the spectra are somewhat different; in particular, the δ-band is less defined in the Phe67Ala80 spectrum, and the 626 nm band is more distinct. Both of these differences suggest that Fe(III)Phe67Ala80 has some high-spin character at pH 7.

Circular Dichroism Spectroscopy

The CD spectrum of Fe(III)Phe67Ala80cyt *c* at pH 7.0 between 190 and 280 nm is very similar to that of native cyt *c* at the same concentration, demonstrating that this mutant has about the same amount of α-helix as native protein (Figure 2.37). The melting temperature of Phe67Ala80 is 84 °C; only 1 degree less than WT and 3 degrees greater than Fe(III)Ala80 (Figure 2.38).

Figure 2.36 UV/vis absorption spectra of (A) h. h. Fe(III)Phe67Ala80cyt *c* at pH 4 (dashed line), 7 (solid line), and 10 (dotted line); and (B) h. h. Fe(II)Phe67Ala80cyt *c* in the presence of O_2 (solid line) and CO (dashed line). Spectra were taken on samples 4–8 μM in protein in 25 mM NaP_i (pH 7.0) in a 1-cm cell.


A
Fe(III)Phe67Ala80
pH 4.0
pH 7.0
pH 10.0
x10
ε (mM-1 cm-1)
λ (nm)
0
20
40
60
80
100
120
140
300
400
500
600
700

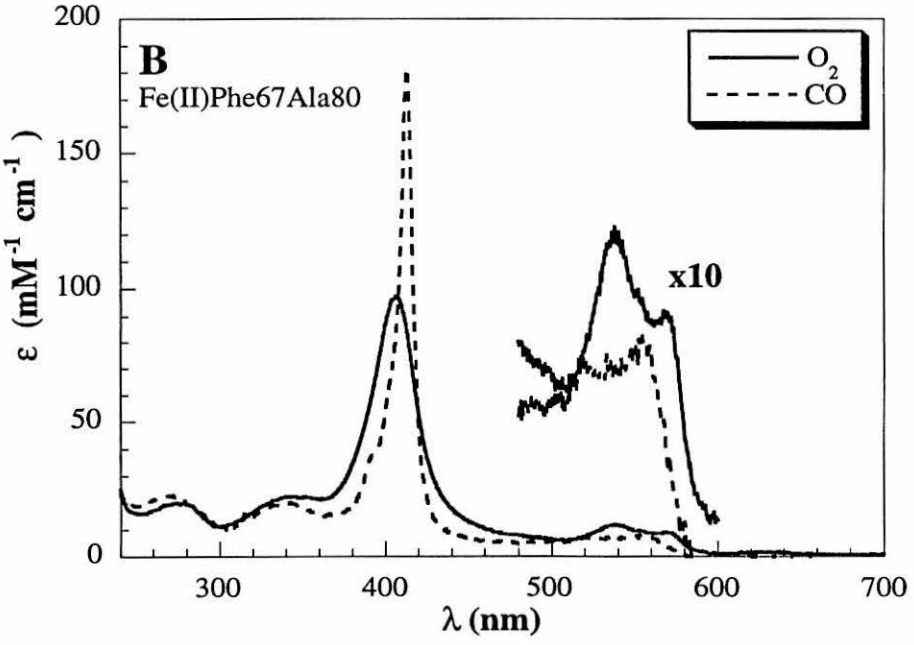

B
Fe(II)Phe67Ala80
O2
CO
x10
ε (mM-1 cm-1)
λ (nm)
0
50
100
150
200
300
400
500
600
700

Figure 2.37 Far-UV CD spectrum of h. h. Phe67Ala80cyt *c*. The spectrum was taken at room temperature on a sample 12 μM in protein in 25 mM NaP_i (pH 7.0) in a 0.1-cm cell.

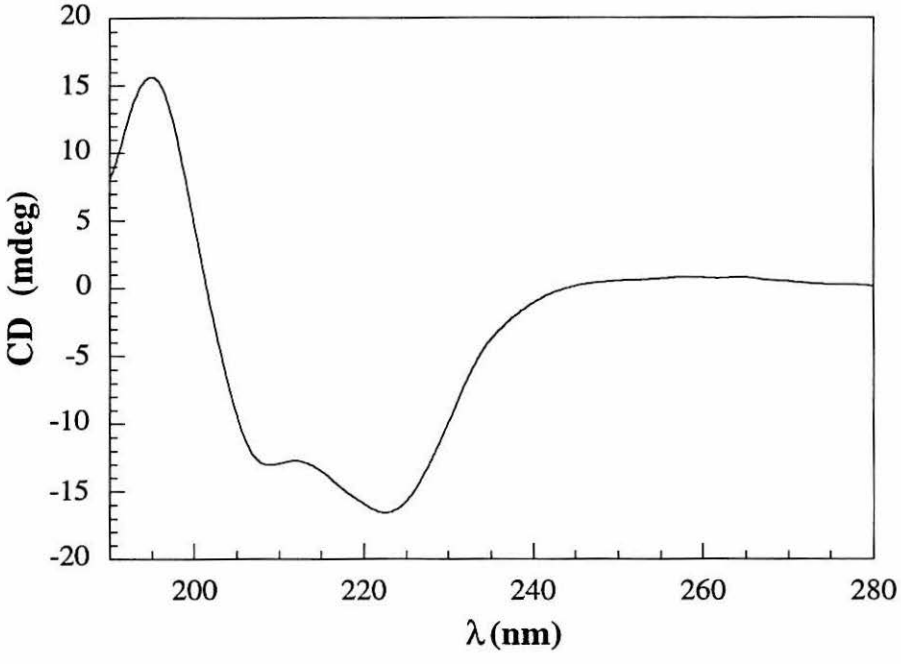
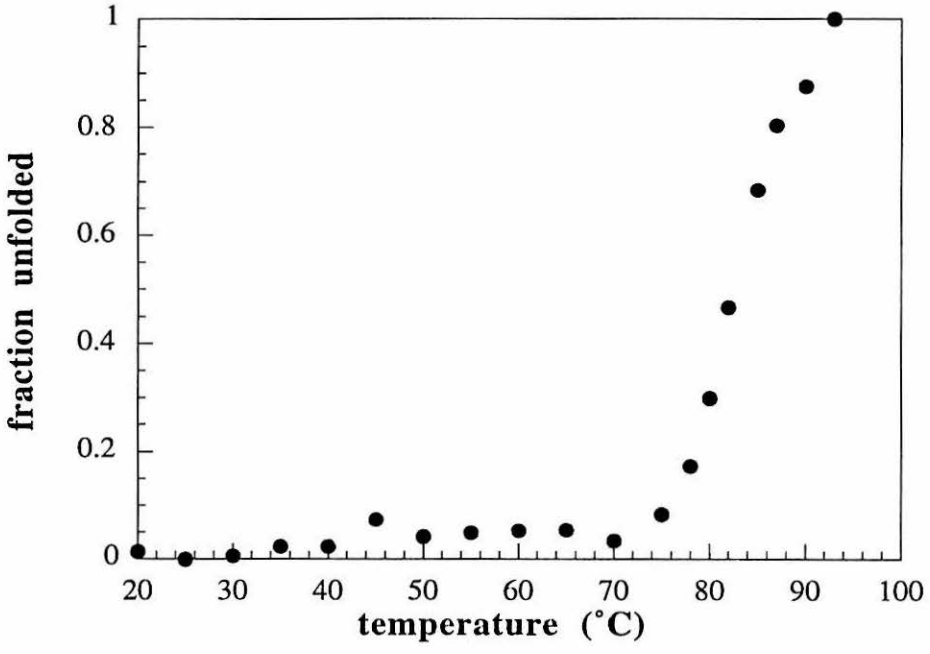

20
15
10
5
0
-5
-10
-15
-20
CD (mdeg)
200
220
240
260
280
λ(nm)

Figure 2.38 CD melting curve for h. h. Fe(III)Phe67Ala80cyt *c*. The data are a plot of the percent of unfolded protein as determined by the temperature dependence of the CD at 222 nm. The sample was 12 μM in protein in 25 mM NaP_i (pH 7.0) in a jacketed 0.1-cm cell.


fraction unfolded
1
0.8
0.6
0.4
0.2
0
20
30
40
50
60
70
80
90
100
temperature (°C)

Ligand Binding to Horse Heart Phe67Ala80cyt *c*

Absorption Spectroscopy

Like h. h. Ala80, h. h. Phe67Ala80 binds a number of small ligands in both the ferric and ferrous oxidation states to give spectra similar to those of Mb. UV/vis absorption data for Phe67Ala80 are presented in Table 2.7. A significant difference exists between the spectra of Fe(III)Phe67Ala80 and of Fe(III)Ala80 without added exogenous ligand; the absorption spectra of Fe(III)Phe67Ala80 are more suggestive of a

Table 2.7

UV/vis absorption data for derivatives of horse heart Phe67Ala80cyt c

Protein	α	β	γ	δ	U. V.
ferrous derivatives:					
deoxyPhe67Ala80	548[a]		406/434[b]	*	282
	8.6		*80.0*[c]		*22.1*
O_2-Phe67Ala80	568	538	406	343	278
	9.2	*12.3*	*97.8*	*22.8*	*20.1*
CO-Phe67Ala80	550–560; broad		413	340	269
	~ 8		*181.7*	*21.0*	*23.7*
ferric derivatives:					
pH 10.0	529		405	356	282
	9.2		*114.0*	*25.0*	*18.4*
pH 7.0	626	522	404	353[b]	274
	2.3	*9.5*	*107.0*	*28*	*21.6*
pH 4.0	627	493	401	*	273
	3.4	*9.1*	*120.8*		*22.4*
CN-Phe67Ala80	542		414	357	273
	10.5		*99.5*	*28.9*	*19.4*

[a]Upright figures indicate wavelength in nm, italics indicate millimolar extinction coefficient.

[b]shoulders.

[c]ε at 406 nm.

* not resolved.

high-spin heme. At low pH, while the Fe(III)Phe67Ala80 Soret band blue-shifts as is expected for the low-spin to high-spin transition, its extinction coefficient only increases to 120.8 $mM^{-1}cm^{-1}$, compared to 164.4 $mM^{-1}cm^{-1}$ for H_2O-Fe(III)Ala80 (Figure 2.36A). One possible explanation for this difference would be that Fe(III)Phe67Ala80 is pentacoordinate at low pH. Heme coordination number has been changed from 6 to 5 in Mb by substituting Val for the distal His in human Mb.[41] Also, *Aplysia* Mb, which has a Leu in the E7 position is pentacoordinate at and below pH 7.[60] However, pentacoordinate high-spin heme proteins typically have Soret bands with an extinction coefficients near or below 100 $mM^{-1}cm^{-1}$ with a significantly larger linewidth than is observed for h. h. Fe(III)Phe67Ala80.[38,41]

Paramagnetic Nuclear Magnetic Resonance Spectroscopy

The ^1H NMR spectrum of Fe(III)Phe67Ala80cyt *c* at pH 7 is characteristic of a low-spin ferric heme, with a small amount of high-spin component as indicated by the downfield peaks at 59 and 55 ppm (Figure 2.39B). This high-spin component vanishes at pH 7.6 (Figure 2.39A). The downfield peaks increase in intensity as pH is lowered; at pH 4.5 the high-spin species dominates (Figure 2.39C). The spectrum of h. h. Fe(III)Phe67Ala80 at low pH differs from that of h. h. Fe(III)Ala80, although it has analogous features with similar shifts and linewidths.

Ligand Photolysis

The bimolecular rate constant of CO recombination to h. h. Fe(II)Phe67Ala80 was measured to be 2.4(5) x 10^4 $M^{-1}s^{-1}$ (Figure 2.40), a value about three times greater than that measured for h. h. Fe(II)Ala80. In Mb mutants (Table 2.4) individual residues in the heme pocket have been shown to regulate rates of ligand-binding to the heme. Here, it has been found that the substituting Tyr67 with Phe increases the CO binding rate.

Figure 2.39 Downfield region of the 500 MHz ^{1}H NMR spectrum of h. h. Fe(III)Phe67Ala80cyt *c* at (A) pH 7.6, (B) pH 7.0, (C) pH 4.5 (pH values not corrected for the isotope effect). Data were collected at 300 K on a 0.25 mM sample in 25 mM NaP_i in D_2O. The residual HDO peak was suppressed with presaturation.

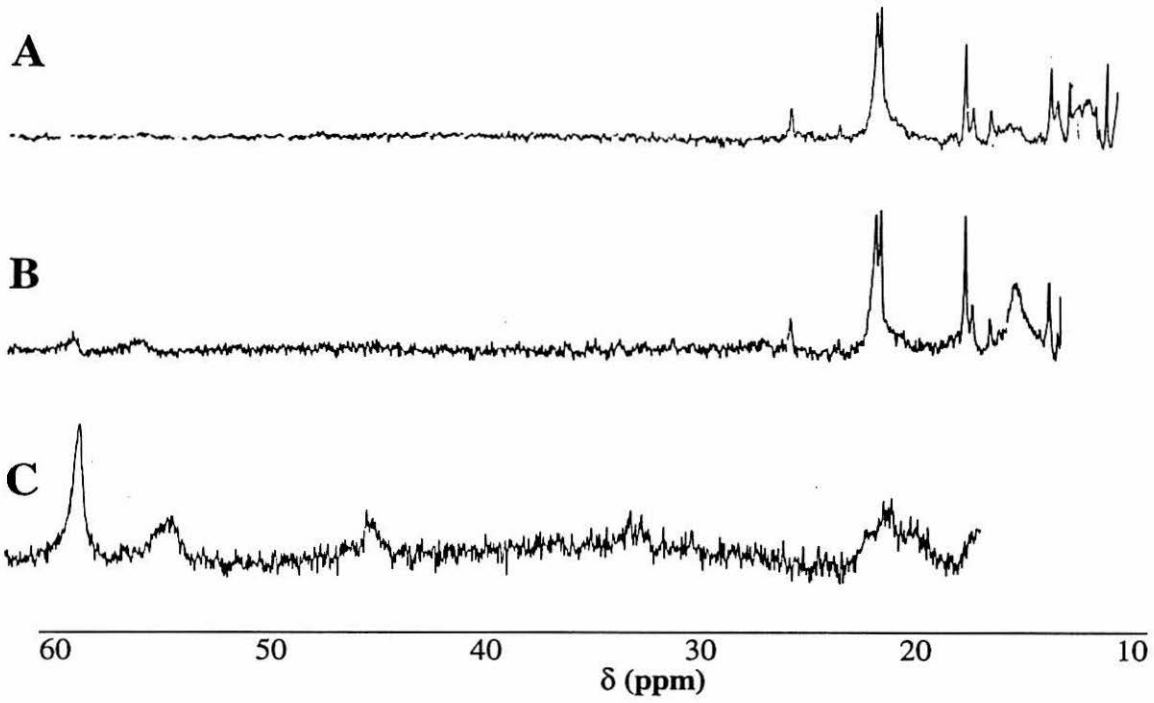

A
B
C
60
50
40
30
20
10
δ (ppm)

Figure 2.40 Transient absorption kinetics for the recombination of CO to h. h. Fe(II)Phe67Ala80cyt *c* after photolysis. The bleach corresponds to the decrease in absorption of the Soret band of the CO-bound species as CO dissociates and the OD increase to the recovery of this absorption as CO binds. The smooth line indicates the fit to a single exponential function under pseudo-first-order conditions. For this experiment, [CO] = 1 mM and $k_{ps} = 26\ s^{-1}$.

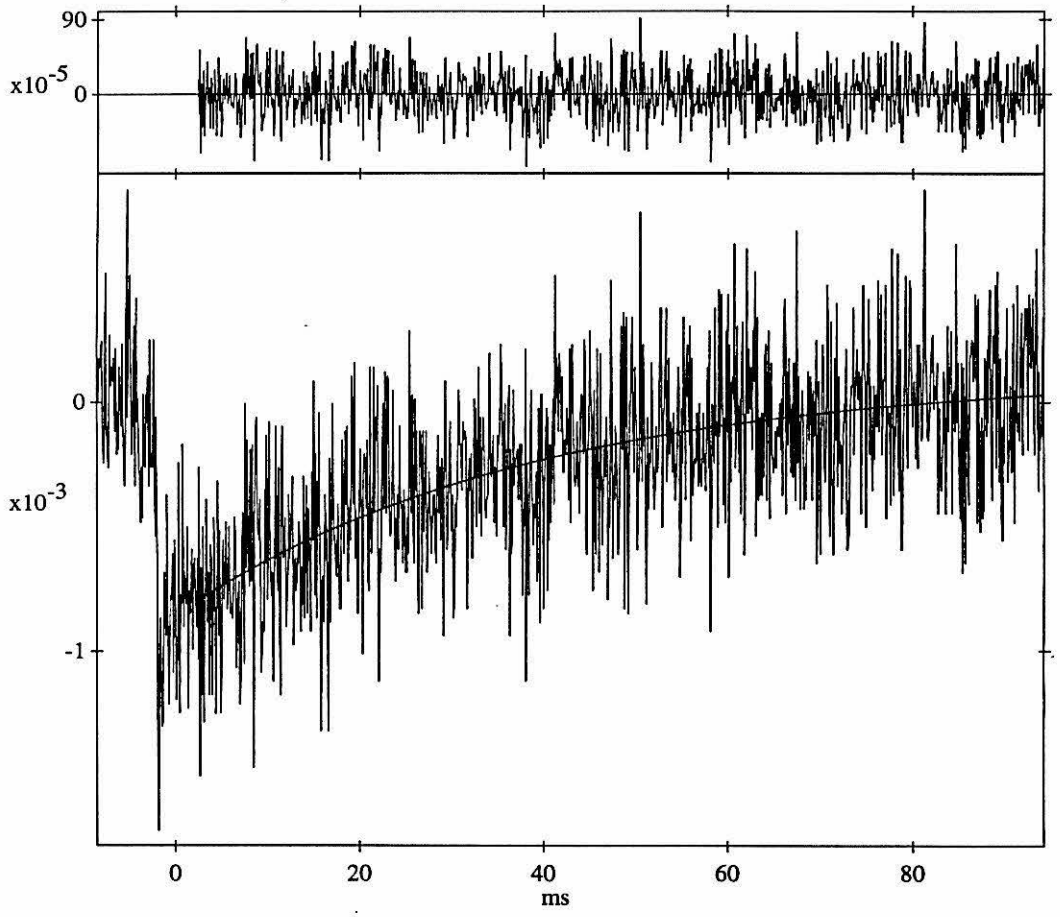
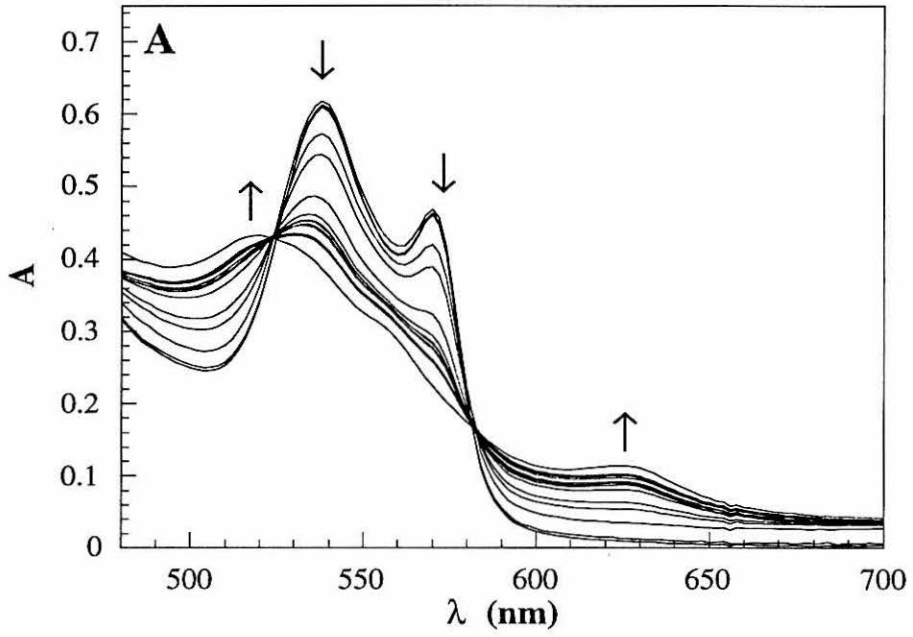
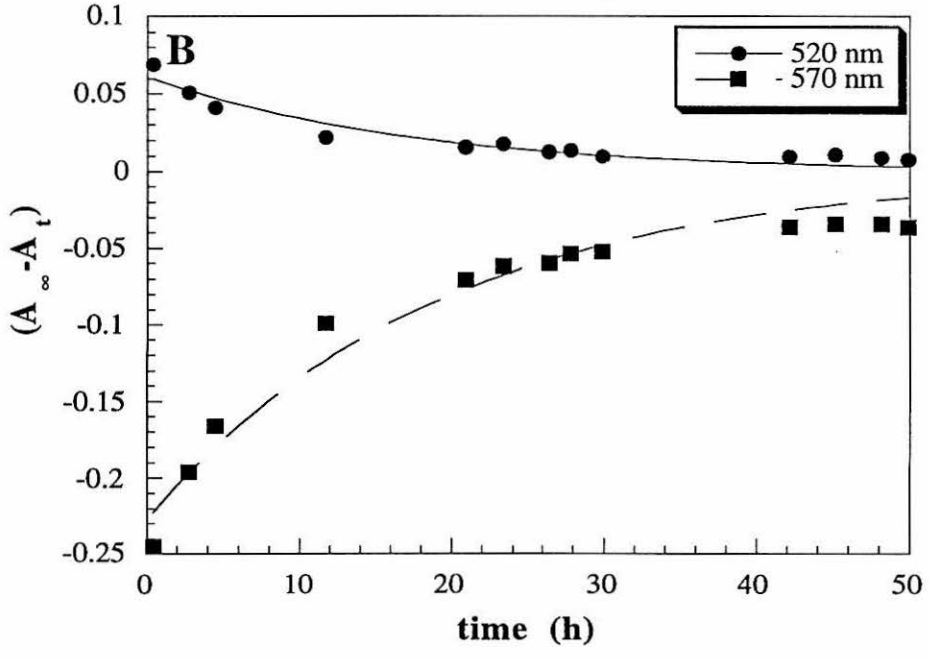

90
x10-5
0
0
x10-3
-1
0
20
40
60
80
ms

Figure 2.41 (A) UV/vis absorption spectra taken over the course of the autoxidation of O_2-Fe(II)Phe67Ala80cyt *c*. Data were recorded on 50 μM protein in 25 mM NaP_i (pH 7.0) at 22 °C. The arrows indicate the spectral changes as the sample autoxidizes. (B) Plot of autoxidation data. The quantity $(A_\infty - A_t)$ is plotted against time and the data are fit to a single exponential to obtain the autoxidation rate, 0.050(6)h^{-1}.


A
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
A
500
550
600
650
700
λ (nm)


B
520 nm
- 570 nm
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
-0.25
(A ∞ -A t)
0
10
20
30
40
50
time (h)

Autoxidation

The rate constant determined for autoxidation of semisynthetic O_2-Fe(II)Phe67Ala80 is 0.050(5) h^{-1}. The kinetics plot and UV/vis data are presented in Figure 2.41. Unlike h. h. O_2-Fe(II)Ala80, the autoxidation proceeds cleanly for this mutant with maintenance of isosbestic points throughout the reaction. The fact that a clean autoxidation reaction is seen when Tyr67 is replaced with Phe provides support for the theory that Tyr67 radical species participate in the autoxidation of h. h. O_2-Fe(II)Ala80.

It is somewhat surprising that Fe(II)Phe67Ala80 forms a kinetically stable dioxygen-bound species as does Fe(II)Ala80. This is because Tyr67 is proposed to hydrogen-bond with the O_2-Fe(II) unit to stabilize this species in Fe(II)Ala80 and this hydrogen bond must be disrupted in the process of autoxidation.[18] The positive edges of the Phe67 and Phe82 rings, however, may play a role in stabilizing O_2-Fe(II)Phe67Ala80. Because thermodynamic and kinetic stabilities of oxy-hemes are related and tend to shown similar trends, it is likely that Phe67Ala80 forms a dioxygen-bound species that is thermodynamically as well as kinetically stable.

Conclusions

Horse heart Ala80cyt *c* and Phe67Ala80cyt *c* have been prepared by semisynthesis in amounts sufficient for extensive biophysical characterization. SDS/PAGE electrophoresis indicates that products with molecular weights similar to that of native cyt *c* were formed, and CD spectroscopy confirms that both proteins contain virtually the same amount of α-helical structure as native cyt *c*. The temperature dependence of the CD spectra shows that the thermal stability of the cytochrome fold is only slightly decreased by the Met80Ala and the Tyr67Phe/Met80Ala mutations.

Both mutants have been shown to bind a variety of axial ligands (O_2, CO, HO^-, H_2O, CN^-). The absorption spectra of axially ligated derivatives of h. h. Ala80 and

Tyr67Ala80 are remarkably similar to analogous myoglobin species. An exception is that the ferrous-deoxy forms of Ala80 and Tyr67Ala80 have Soret maxima significantly blue-shifted from the observed maximum in Mb, with a shoulder at lower energy, suggesting the presence of a mixture of spin states. MCD spectroscopy indicates that deoxy-Fe(II)Ala80cyt *c* is indeed part high-spin and part low-spin. The oxidized form of Ala80 has also been investigated by MCD; the data indicate that the protein is low-spin at pH 7.0 and high-spin with His-H_2O ligation at pH 4.5. The UV/vis spectrum of Fe(III)Ala80 at pH 7 is very similar to that of Mb at high pH which is known to have His-OH^- ligation. The ligand-field parameters calculated from the **g** values for Fe(III)Ala80 support the assignment of His-OH^- ligation to a low-spin heme at neutral pH. The range of hyperfine shifts observed in the ^1H NMR spectra of paramagnetic h. h. Fe(III)Ala80 and Fe(III)Phe67Ala80 also indicate that both mutants are mostly low-spin at pH 7 and high-spin below pH 6.

The dioxygen adduct of h. h. Fe(II)Ala80cyt *c* is remarkably stable. Ala80 binds dioxygen with an affinity greater than that reported for most mammalian Mbs and autoxidizes with a half-life of three days ($k = 0.010(3)\ h^{-1}$) while Mb autoxidizes in hours. The high stability of O_2-Fe(II)Ala80cyt *c* with respect to O_2-Fe(II)Mb is especially surprising in light of reports that nearly all distal pocket mutants of Mb show dramatically decreased O_2 affinities and high autoxidation rates.[16,20] A notable exception is Leu29 → Phe sperm whale Mb, which is extremely stable in the oxy state due to favorable interactions of the iron-bound dioxygen with the positive edge of Phe29.[20a] Molecular modeling suggests that Phe82 could play a similar role in Ala80cyt *c*. In addition, modeling shows that Tyr67 is positioned in the cytochrome pocket to donate a hydroxyl proton to dioxygen, playing a role analogous to the distal histidine in Mb. A Tyr-dioxygen interaction has in fact been identified in *Ascaris* Hb which has a remarkably high dioxygen avidity.[58c] Horse heart Phe67Ala80cyt *c* has a higher autoxidation rate ($k = 0.050(6)\ h\text{-}^1$) than h. h. Ala80, providing experimental indication

that Tyr67 may aid in stabilizing this species in Ala80. However, O_2-Fe(II)Phe67Ala80 is quite stable despite the lack of a potential Tyr67-O_2 interaction.

The kinetics of CO recombination to h. h. Fe(II)Ala80 and Fe(II)Phe67Ala80 after photolysis have been investigated by transient absorption spectroscopy. Carbon monoxide recombines anomalously slowly to both h. h. Fe(II)Ala80 ($k = 8.7(8) \times 10^3\ M^{-1}s^{-1}$) and to h. h. Fe(II)Phe67Ala80 ($k = 2.5(5) \times 10^4\ M^{-1}s^{-1}$) relative to its reaction with h. h. Mb ($k = 5.0 \times 10^5\ M^{-1}s^{-1}$). The quantum yields of CO and O_2 dissociation from Fe(II)Ala80 relative to Mb ($\Phi_{rel,CO} = 0.02$; $\Phi_{rel,O_2} < 0.01$) and from Phe67Ala80 ($\Phi_{rel,CO} = 0.02$) are exceedingly low; geminate recombination was not observed on the nanosecond timescale for either mutant.

The low yields with which O_2 and CO escape from the heme pocket of the mutants in addition to the slow CO recombination demonstrate that the iron site is much less accessible in the cyt *c* variants than in Mb. Dynamics calculations suggest that particular amino acids in Mb facilitate diffusion between the solvent and the heme iron.[52] Such a ligand-binding channel apparently is not available in the cytochrome fold. The striking stability of O_2-Fe(II)Ala80 and of O_2-Fe(II)Phe67Ala80 can be partially attributed to the highly protected ligand-binding site.

References and Notes

(1) For reviews on the structure and function of cytochromes *c*, see (a) Moore, G. R.; Pettigrew, G. W. *Cytochromes c; Evolutionary, Structural and Physicochemical Aspects*; Springer-Verlag: Berlin, 1990. (b) Pettigrew, G. W.; Moore, G. R. *Cytochromes c; Biological Aspects*; Springer-Verlag: Berlin, 1987. (c) Dickerson, R. E.; Timkovich, R. In *The Enzymes*; Boyer, P., Ed.; Academic Press: New York, 1975; Vol. XI, pp 397-547.

(2) Makinen, M. W.; Churg, A. K. In *Iron Porphyrins*; Lever, A. P. B., Gray, H. B., Eds.; Addison Wesley: London, 1983; pp 141-235.

(3) X-ray and NMR structures of cytochromes *c* include: (a) oxidized horse heart: Bushnell, G. W.; Louie, G. V.; Brayer, G. D. *J. Mol. Biol.* **1990,** *214,* 585-595. (b) reduced horse heart: Qi, P. W.; DiStefano, D. L.; Wand, A. J. *Biochemistry*

1994, *33,* 6408-6417. (c) oxidized baker's yeast (*Saccharomyces cerevisiae*) iso-1: Berghuis, A. M.; Brayer, G. D. *J. Mol. Biol.* **1992,** *223,* 959-976. (d) reduced baker's yeast iso-1: Louie, G. V.; Brayer, G. D. *J. Mol. Biol.* **1990,** *214,* 528-555. (e) reduced albacore tuna heart: Takano, R.; Dickerson, R. E. *J. Mol. Biol.* **1981,** *153,* 79-94. (f) oxidized albacore tuna heart: Takano, R.; Dickerson, R. E. *J. Mol. Biol.* **1981,** *153,* 95-115. (g) reduced bonito: Tanaka, N.; Tamane, T.; Tsukihara, T.; Ashida, T.; Kakudo, M. *J. Biochem.* **1975,** *77,* 147-162. (h) oxidized rice (*Oryza sativa*): Ochi, H.; Hata, Y.; Tanaka, N.; Kakudo, M.; Sakurai, T.; Aihara, S.; Morita, Y. *J. Mol. Biol.* **1983,** *166,* 407-418.

(4) Duus, J. Ø.; Low, D. W.; Gray, H. B., in preparation.

(5) For reviews on the semisynthetic method, see: Wallace, C. J. A. *Curr. Opin. Biotech.* **1995,** *6,* 403-410. Humphries, J.; Offord, R. E.; Smith, R. A. G. *Curr. Opin. Biotech.* **1991,** *2,* 539-543. Offord, R. E. *Protein Engineering* **1987,** *1,* 151-157.

(6) (a) Corradin, G.; Harbury, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1971,** *68,* 3036-3039. (b) Corradin, G.; Harbury, H. A. *Biochem. Biophys. Res. Comm.* **1974,** *61,* 1400-1406.

(7) (a) Wallace, C. J. A.; Corradin, G.; Marchiori, F.; Borin, G. *Biopolymers* **1986,** *25,* 2121-2132. (b) Koul, A. K.; Wasserman, G. F.; Warme, P. K. *Biochem. Biophys. Res. Commun.* **1979,** *89,* 1253-1259. (c) Wallace, C. J. A.; Corthésy, B. E. *Protein Eng.* **1986,** *1,* 23-27. (d) Wasserman, G. F.; Nix, P. T.; Koul. A. K.; Warme, P. K. *Biochim. Biophys. Acta* **1980,** *623,* 457-460. (e) ten Kortenaar, P. B. W.; Adams, P. J. H. M.; Tesser, G. I. *Proc. Natl. Acad. Sci. U.S.A.* **1985,** *82,* 8279-8283. (f) Nix, P. T.; Warme, P. K. *Biochim. Biophys. Acta* **1979,** *578,* 413-427. (g) Wallace, C. J. A.; Offord, R. E. *Biochem. J.* **1979,** *179,* 169-182. (h) Boon. P. J.; van Raay, A. J. M.; Tesser, G. I.; Nivard, R. J. F. *FEBS Lett.* **1979,** *108,* 131-145.

(8) Merrifield, R. B. *Adv. Enzymol.* **1969,** *32,* 221-296. Kent, S. B. H. *Ann. Rev. Biochem.* **1988,** *57,* 957-989.

(9) (a) Barstow, L. E.; Young, R. S.; Yakali, E.; Sharp, J. J.; O'Brien, J. C.; Berman, P. W.; Harbury, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1977,** *74,* 4248-4250. (b) Wallace, C. J. A.; Mascagni. P.; Chait, B. T.; Collawn, J. F.; Paterson, Y.; Proudfoot, A. E. I.; Kent, S. B. H. *J. Biol. Chem.* **1989,** *264,* 15199-15209. (c) Raphael, A. L.; Gray, H. B. *Proteins* **1989,** *6,* 338-340. (d) Raphael, A. L.; Gray, H. B. *J. Am. Chem. Soc.* **1991,** *113,* 1038-1040. (e) Wuttke, D. S.; Bjerrum, M. J.; Winkler, J. R.; Gray, H. B. *Science* **1992,** *256,* 1007-1009. (f) Wuttke, D. S.;

Bjerrum, M. J.; Winkler, J. R.; Gray, H. B. *Biochim. Biophys. Acta* **1992,** *1101,* 168-170. (g) Wallace, C. J. A.; Clark-Lewis, I. *J. Biol. Chem.* **1992,** *267,* 3852-3861. (h) Frauenhoff, M. M.; Scott, R. A. *Proteins* **1992,** *14,* 202-212. (i) Wuttke, D. S.; Gray, H. B.; Fisher, S. L.; Imperiali, B. *J. Am. Chem. Soc.* **1993,** *115,* 8455-8456. (j) Ueda, T.; Kimura, S.; Imanishi, Y. *J. Chem. Soc. P1* **1994,** *2,* 219-224.

(10) Boswell, A. P.; Moore, G. R.; Williams, R. J. P.; Wallace, C. J. A.; Boon, P. J.; Nivard, R. J. F.; Tesser, G. I. *Biochem. J.* **1981,** *193,* 493-502.

(11) Lu, Y.; Casimiro, D. R.; Bren, K. L.; Richards, J. H.; Gray, H. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993,** *90,* 11456-11459. The characterization of this non-functional mutant of yeast iso-1-cyt *c* is also presented in Chapter 3 of this thesis.

(12) The ligand-binding properties and structure of myoglobin are reviewed in: Antonini, E.; Brunori, M. *Hemoglobin and Myoglobin in Their Reactions with Ligands*; North-Holland: Amsterdam, 1971.

(13) Phillips, S. E. V. *J. Mol. Biol.* **1980,** *142,* 531-554.

(14) La Mar, G. N.; Hauksson, J. B.; Dugad, L. B.; Liddell, P. A.; Venkataramana, N.; Smith, K. M. *J. Am. Chem. Soc.* **1991,** *113,* 1544-1550.

(15) For a review on dioxygen storage by natural and synthetic systems, see: Jameson, G. E.; Ibers, J. A. In *Bioinorganic Chemistry*; Bertini, I., Gray, H. B., Lippard, S. J., Valentine, J. S., Eds.; University Science Books: Mill Valley, California, 1994; pp 167-252.

(16) The mechanisms of ligand recognition in myoglobin are reviewed in: Springer, B. A.; Sligar, S. G.; Olson, J. S.; Phillips, G. N., Jr. *Chem. Rev.* **1994,** *94,* 699-714.

(17) Balch, A. L.; Chan, Y. W.; Cheng, R. J.; La Mar, G. N.; Latosgrazynski, L.; Renner, M. W. *J. Am. Chem. Soc.* **1984,** *106,* 7779-7785.

(18) Brantley, R. E., Jr.; Smerdon, S. J.; Wilkinson, A. J.; Singleton, E. W.; Olson, J. S. *J. Biol. Chem.* **1993,** *268,* 6995-7010.

(19) Philips, S. E. V.; Schoenborn, B. P. *Nature* **1981,** *292,* 81-82. Pauling, L. *Nature* **1964,** *203,* 182-183.

(20) (a) Carver, T. E.; Brantley, R. E., Jr.; Singleton, E. W.; Arduini, R. M.; Quillin, M. L.; Phillips, G. N., Jr.; Olson, J. S. *J. Biol. Chem.* **1992,** *267,* 14443-14450. (b) Rohlfs, R. J.; Mathews, A. J.; Carver, T. E.; Olson, J. S.; Springer, B. A.; Egeberg, K. D.; Sligar, S. G. *J. Biol. Chem.* **1990,** *265,* 3168-3176. (c) Springer, B. A.; Egeberg, K. D.; Sligar, S. G; Rohlfs, R. J.; Mathews, A. J.; Olson, J. S. *J. Biol. Chem.* **1989,** *264,* 3057-3060. (d) Olson, J. S.; Mathews, A. J.; Rohlfs, R. J.;

Springer, B. A.; Egeberg, K. D.; Sligar, S. G.; Tame, J.; Renaud, J.-P.; Nagai, K. *Nature* **1988,** *336,* 265-266.

(21) Frydman, R. B.; Frydman, B. *Acc. Chem. Res.* **1987,** *20,* 250-256.

(22) Norvell, J. C.; Nunes, A. C.; Schoenborn, B. P. *Science,* **1975,** *190,* 568-569. Heidner, E. J.; Ladner, R. C.; Perutz, M. F. *J. Mol. Biol.* **1976,** *104,* 707-722.

(23) Gibson, Q. H. *Biochem. Soc. Trans.* **1990,** *18,* 1-6. Gibson, Q. H. *J. Biol. Chem.* **1989,** *264,* 20155-20158. Findsen, E. W.; Ondrias, M. R. *Photochem. Photobiol.* **1990,** *51,* 741-748.

(24) Petrich, J. W.; Martin, J. L. *Chem. Phys.* **1989,** *131,* 31-47.

(25) Austin, R. H.; Beeson, K. W.; Eisenstein, L.; Frauenfelder, H.; Gunsalus, C. *Biochemistry* **1975,** *14,* 5355-5373. Ansari, A.; DiIorio, E. E.; Dlott, D. D.; Frauenfelder, H.; Iben, I. E. T.; Langer, P.; Roder, H.; Sauke, T. B.; Shyamsunder, E. *Biochemistry* **1986,** *25,* 3139-3146. Henry, E. R.; Sommer, J. H.; Hofrichter, J.; Eaton, W. A. *J. Mol. Biol.* **1983,** *166,* 443-451.

(26) Ansari, A.; Jones, C. M.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Biochemistry* **1994,** *33,* 5128-5145. Lim, M. H.; Jackson, T. A.; Anfinrud, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993,** *90,* 5801-5804. Neinhaus, G. U.; Mourant, J. R.; Frauenfelder, H. *Proc. Natl. Acad. Sci. U.S.A.* **1992,** *89,* 2902-2906. Xie, X.; Simon, J. D. *Biochemistry* **1991,** *30,* 3682-3692. Petrich, J. W.; Lambry, J.-C.; Kuczera, K.; Karplus, M.; Poyart, C.; Martin, J.-L. *Biochemistry,* **1991,** *30,* 3975-3987. Jongeward, K. A.; Magde, D.; Taube, D. J.; Marsters, J. C.; Traylor, T. G.; Sharma, V. S. *J. Am. Chem. Soc.* **1988,** *110,* 380-387.

(27) Gibson, Q. H.; Regan, R.; Elber, R.; Olson, J. S.; Carver, T. E. *J. Biol. Chem.* **1992,** *267,* 22022-22034.

(28) Kirschner, S. *Inorg. Syn.* **1957,** *5,* 186-188.

(29) Margoliash, E.; Frohwirt, N. *Biochem. J.* **1959,** *71,* 570-572.

(30) Berry, E. A.; Trumpower, B. L. *Anal. Biochem.* **1987,** *161,* 1-15.

(31) Taylor, C. P. S. *Biochim. Biophys. Acta* **1977,** *491,* 137-149.

(32) Bohan, T. H. *J. Magn. Res.* **1977,** *26,* 109-118.

(33) Blumberg, W. E.; Peisach, J. In *Probes of Structure and Function of Macromolecules and Membranes*; Chance, B., Yonetani, T., Mildvan, A. S., Eds.; Academic Press: New York, 1971; Vol. 2, pp 215-227.

(34) Rux, J. J.; Dawson, J. H. *FEBS Lett.* **1991,** *290,* 49-51.

(35) Giardina, B.; Amiconi, G. *Meth. Enzymol.* **1981,** *76,* 417-427.

(36) Taylor, J. R. *An Introduction to Error Analysis*; University Science: Mill Valley, CA, 1982.

(37) Matsuoka, A.; Kobayashi, N.; Shikama, K. *Eur. J. Biochem.* **1992,** *210,* 337-341.

(38) Bogumil, R.; Maurus, R.; Hildebrand, D. P.; Brayer, G. D.; Mauk, A. G. *Biochemistry* **1995,** *34,* 10483-10490.

(39) The presence of a low-energy (620-630 nm) absorption band is often taken to be indicative of high-spin ferric heme. This band is present in Fe(III)Ala80cyt *c* at pH 7, although it is significantly weaker than at lower pH values.

(40) Palmer, G. In *Iron Porphyrins,* Lever, A. P. B., Gray, H. B., Eds.; Addison-Wesley: London, 1983; Part 2, pp 43-85.

(41) Ikeda-Saito, M.; Hori, H.; Andersson, L. A.; Prince, R. C.; Pickering, I. J.; George, G. N.; Sanders, C. R., II; Lutz, R. S.; McKelvey, E. J.; Mattera, R. *J. Biol. Chem.* **1992,** *267,* 22843-22852.

(42) Salmeen, I.; Palmer, G. *J. Chem. Phys.* **1968,** *48,* 2049-2053.

(43) Blum, H.; Harmon, H. J.; Leigh, J. S.; Salerno, J. C.; Chance, B. *Biochim. Biophys. Acta* **1978,** *502,* 1-10.

(44) Ikeda, M.; Iizuka, T.; Takao, H.; Hagihara, B. *Biochim. Biophys. Acta* **1974,** *336,* 15-24.

(45) For reviews on MCD, see: (a) Cheesman, M. R.; Greenwood, C.; Thomson, A. J. *Adv. Inorg. Chem.* **1991,** *36,* 201-255. (b) Holmquist, B. *Adv. Enzymol.* **1986,** *130,* 270-290. (c) Myer, Y. P. *Curr. Top. Bioenerg.* **1985,** *14,* 149-188. (d) Dawson, J. H.; Dooley, D. M. In *Iron Porphyrins*; Lever, A. P. B., Gray, H. B., Eds.; VCH: New York, 1989; pp 1-135.

(46) Bertini, I.; Luchinat, C. *NMR of Paramagnetic Molecules in Biological Systems*; Benjamin/Cummings: Menlo Park, 1986.

(47) La Mar, G. N.; Walker, F. A. In *The Porphyrins*; Dolphin, D., Ed.; Academic Press: New York, 1979; pp 61-157.

(48) Senn, H.; Böhme, H.; Wüthrich, K. *Biochim. Biophys. Acta* **1984,** *789,* 311-323, and references therein.

(49) Turner, D. L. *Eur. J. Biochem.* **1995,** *227,* 829-837.

(50) Gibson, Q. H.; Olson, J. S.; McKinnie, R. E.; Rohlfs, R. J. *J. Biol. Chem.* **1986,** *261,* 10228-10239.

(51) Rohlfs, R. J.; Olson, J. S.; Gibson, Q. H. *J. Biol. Chem.* **1988,** *263,* 1803-1813.

(52) Case, D. A.; Karplus, M. *J. Mol. Biol.* **1979,** *132,* 343-368. Elber, R.; Karplus, M. *J. Am. Chem. Soc.* **1990,** *112,* 9161-9175. Kottalam, J.; Case, D. A. *J. Am. Chem. Soc.* **1988,** *110,* 7690-7697.

(53) Colón, J. L., unpublished results.

(54) King, N. K.; Winfield, M. E. *J. Biol. Chem.* **1963,** *238,* 1520-1528.

(55) Egeberg, K. D.; Springer, B. A.; Sligar, S. G.; Carver, T. E.; Rohlfs, R. J.; Olson, J. S. *J. Biol. Chem.* **1990,** *265,* 11788-11795.

(56) Bellelli, A.; Blackmore, R. S.; Gibson, Q. H. *J. Biol. Chem.* **1990,** *265,* 13595-13600.

(57) Brown, W. D.; Mebine, L. B. *J. Biol. Chem.* **1969,** *244,* 6696-6701.

(58) (a) Qin, J.; La Mar, G. N.; Ascoli, F.; Bolognesi, M.; Brunori, M. *J. Mol. Biol.* **1992,** *224,* 891-897. (b) Vyas, K.; Rajarathnam, K.; Yu, L. P.; Emerson, D.; La Mar, G. N.; Krishnamoorthi, R.; Mizukami, H. *J. Biol. Chem.* **1993,** *268,* 14826-14835. (c) De Baere, I.; Perutz, M. F.; Kiger, L.; Marden, M. C.; Poyart, C. *Proc. Natl. Acad. Sci. U.S.A.* **1994,** *91,* 1594-1597.

(59) Bren, K. L.; Wuttke, D. S. unpublished results.

(60) Bolognesi, M.; Onesti, S.; Gatti, G.; Coda, A.; Ascenzi, P.; Brunori, M. *J. Mol. Biol.* **1989,** *205,* 529-544.

Chapter 3

Expression, Characterization, and Ligand-Binding Properties of a *Saccharomyces cerevisiae* Ala80-iso-1-cytochrome *c* Variant[1]

Background

Protein engineering has proven to be a remarkably useful tool in studies of structure/function relationships. Of particular utility to the bioinorganic chemist is the ability to alter metal ligation in metalloproteins through site-directed mutagenesis. Engineering metalloprotein active sites has shed much light on the mechanisms proteins utilize to tune metal cofactor reactivity. For example, altering the identity and character of the proximal His in C*c*P has led to a greater understanding of the mechanisms of action of peroxidases and other heme-containing redox enzymes.[2] The substitution of the copper-ligating Cys112 with Asp in the blue-copper protein azurin lent experimental proof to the theory that the intense blue color of azurin is due to a Cys $\rightarrow$ Cu(II) charge-transfer transition;[3] in addition, this mutation greatly alters the electron-transfer activity of azurin.[4] As discussed in this thesis and elsewhere,[5-7] substituting the cyt *c* heme axial ligand Met80 with other ligating and non-ligating amino acids has provided valuable information on the manner in which ligands modulate heme reactivity and also has allowed the use of cyt *c* as a model for other heme proteins such as cytochrome b_5,[5] cytochrome P-450[5] and Mb.[7]

The preparation of cyts *c* with altered heme ligation has until recently been limited to the technique of semisynthesis.[8] Although semisynthesis has proven to be practical for preparing a variety of cyt *c* mutants, only residues in the portion of the protein corresponding to residues 66–104 may be altered. In addition, the amount of protein produced is limited by the amount of 39-mer peptide that can be synthesized in pure form.[9] This makes the preparation of large (> 100 mg) amounts of protein laborious and expensive, for the Ala80cyt *c* variant in particular due to small reconstitution yields. The total synthesis of cyt *c* has been reported, but the small yields of the peptide synthesis, heme attachment, and protein refolding steps preclude the production of multi-milligram amounts of protein.[10] Therefore, to generate significant quantities of product, it is attractive to produce Ala80cyt *c* by recombinant means.

A great number of cyt *c* mutants have been produced by recombinant methodologies.[11] The expression of cyts *c* must be done in eukaryotic organisms that contain the mechanism for the covalent attachment of the heme to the polypeptide post-translationally. The most favored host is baker's yeast (*Saccharomyces cerevisiae, S. c.*)[12] Smith and coworkers have utilized a multi-copy plasmid (YEp) for overexpressing *S. c.* iso-1-cyt *c* (sequence in Figure 3.1) under the control of its natural promoter.[13] The host cells are grown in nonfermentable carbon media so that a functional gene product is required for aerobic respiration and cell growth. Therefore, cyt *c* mutants with reduction potentials significantly different from native cannot be expressed by this method because they will not function properly in the mitochondrial respiratory chain, thus excluding replacements of the heme axial ligands (His18 and Met80),[14] with the exception of the Arg18 mutant which has the same reduction potential as wild-type.[15] A means around this limitation is to utilize a dual-gene vector bearing both the target nonfunctional mutant (*cyc1-*) as well as a functional cyt *c* gene (*cyc1+*) to support growth of the cells in a medium containing non-fermentable carbon (Figure 3.2).[16] In order to facilitate separation of the two gene products, a di-His chelating site designed by Arnold and coworkers[17] was incorporated into the functional cyt *c* gene so that the two isoforms may be separated by metal-affinity chromatography (Figure 3.3).[16] This methodology allowed the preparation of Ala80cyt *c* in quantities large enough for detailed biophysical characterization, including an NMR study. In this chapter, the initial characterization of this mutant is described. A detailed NMR study of *S. c.* Ala80cyt *c* is presented in the following chapters.

Materials and Methods

General

Water used for solutions was either doubly distilled and passed through deionizing and organic removal cartridges on a Corning Mega-Pure system or purified

Figure 3.1 The amino acid sequence of *S. c.* iso-1-cyt *c*. The axial ligands (positions 18 and 80) are in bold type and the Cys thioether linkages (positions 14 and 17) are underlined. The sequence begins at residue -5 because the numbering system is based on mammalian sequences. Tml is trimethyllysine.

-5						Ac-Thr		Glu	Phe	Lys	Ala
1	Gly	Ser	Ala	Lys	Lys	Gly	Ala	Thr	Leu	Phe
11	Lys	Thr	Arg	Cys	Leu	Gln	Cys	**His**	Thr	Val
21	Glu	Lys	Gly	Gly	Pro	His	Lys	Val	Gly	Pro
31	Asn	Leu	His	Gly	Ile	Phe	Gly	Arg	Gln	Ser
41	Gly	Gln	Ala	Glu	Gly	Tyr	Ser	Tyr	Thr	Asp
51	Ala	Asn	Ile	Lys	Lys	Asn	Val	Leu	Trp	Asp
61	Gly	Asn	Asn	Met	Ser	Glu	Tyr	Leu	Thr	Asn
71	Pro	Tml	Lys	Tyr	Ile	Pro	Gly	Thr	Lys	**Met**
81	Ala	Phe	Gly	Gly	Leu	Lys	Lys	Glu	Lys	Asp
91	Arg	Asn	Asp	Leu	Ile	Thr	Tyr	Leu	Lys	Lys
101	Ala	Cys	Gly-OH							

Figure 3.2 Vector construct for expression of *S. c.* Met80Ala (*cyc1-*) and Leu58His (*cyc1+*) iso-1-cytochrome *c*. The shaded boxes contain the promoter and upstream activation sequences for cyc1 transcription. *Leu+* and ampicillin resistance markers are present for plasmid selection in yeast and *E. coli*, respectively.

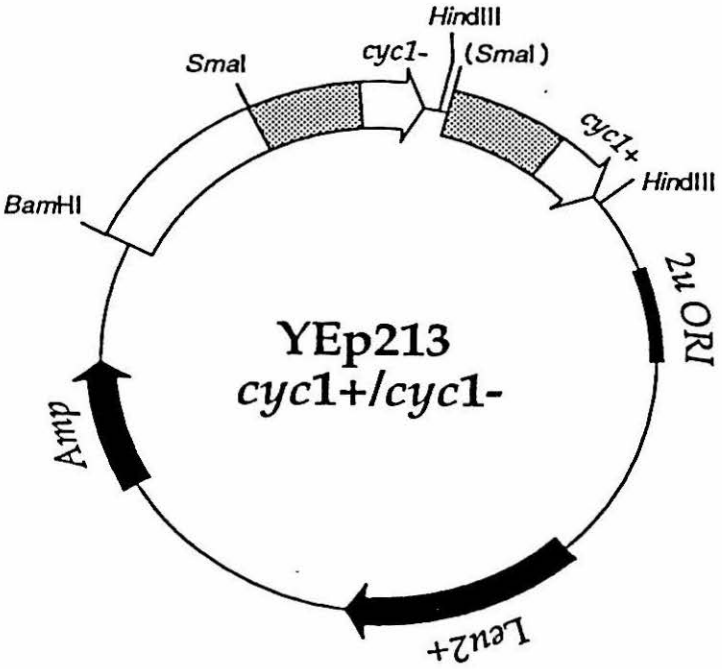
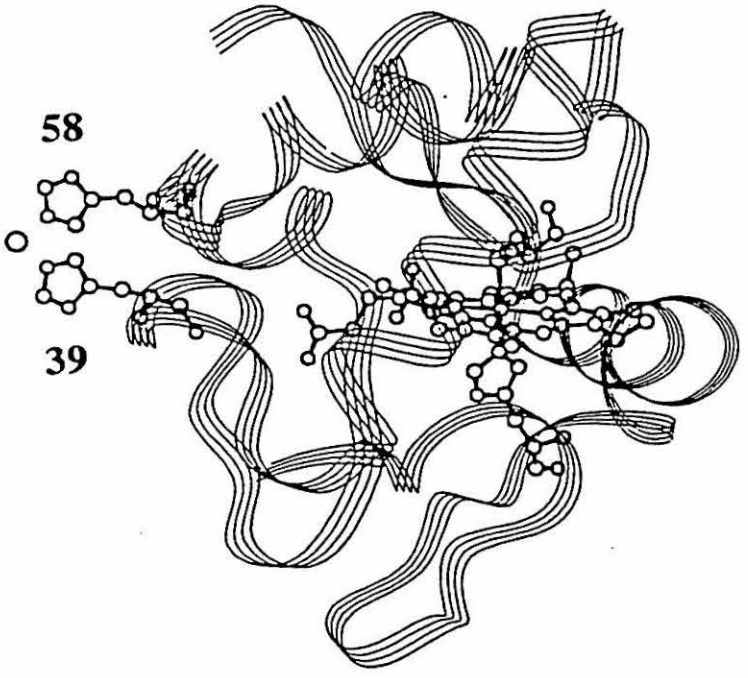

HindIII
SmaI
cyc1-
(SmaI)
cyc1+
HindIII
BamHI
2μ ORI
YEp213
cyc1+/cyc1-
Amp
Leu2+

Figure 3.3 Molecular model of *S. c.* Leu58His cytochrome *c*, indicating the location of the surface di-histidine site for high-affinity binding to immobilized metal complexes. Energy minimization calculations were performed on the 1.23-Å structure of yeast iso-1-cyt c^{24} using BIOGRAF, version 3.0.


58
39

using a Barnstead Nanopure purification system. Bacto-yeast extract, bacto-peptone, bacto-tryptone, bacto-agar, and yeast nitrogen base without amino acids were obtained from Difco. *Hin*dIII was obtained from Gibco. Taq polymerase, NTP, T4 DNA ligase, T4 DNA polymerase, and X-gal were purchased from Boehringer Mannheim. Glycerol was of reagent grade. All other reagents were of analytical grade. Oligonucleotides were synthesized at the Caltech Biopolymer Synthesis Facility under the direction of Susannah Horvath and purified by ethanol precipitation followed by washing with ice-cold 70% ethanol. All oligonucleotides and DNA samples were stored frozen at -20 °C. Centrifugation was performed in a Sorvall Superspeed RC-2B refrigerated centrifuge. The FPLC apparatus used for protein purification and the transient absorption system are described in Chapter 2. Protein concentration and storage were performed as described in Chapter 2. YPD is 1% (w/v) bacto-yeast extract, 2% bacto-peptone, 2% dextrose. YPG is 1% bacto-yeast extract, 2% bacto-peptone, 3% (v/v) glycerol. SD is 0.67% yeast nitrogen base without amino acids, 2% dextrose, 0.002% histidine, 0.002% tryptophan. LB is 1% bacto-tryptone, 0.5% bacto-yeast extract, 1% NaCl. SOC is 2% bacto-tryptone, 0.5% bacto-yeast extract, 0.05% NaCl, and 0.0186% KCl plus 5 mL 2 M $MgCl_2$ and 20 mL 1 M dextrose. Culture plates were made by adding 2% agar to the desired media. TE is 10 mM Tris, 1 mM EDTA. TAE is 40 mM Tris-acetate, 1 mM EDTA.

Site-directed Mutagenesis and Molecular Cloning

The Met80Ala mutation in the yeast iso-1-cyt *c* (*cyc1*) gene was introduced by the procedure developed by Kunkel[18] using the Muta-Gene kit from Bio-Rad customized for these experiments. This procedure is described in detail elsewhere.[19] We found that a more efficient procedure to introduce mutations was by the polymerase chain reaction (PCR) method which is described here. This was used to mutate position 39 from Gln back to the native His on the *cyc1* gene (Appendix A) in which Cys102 had already been

replaced with Ser.[20] In addition, the Met80Ala mutant with His39 in place of Gln was produced. The mutagenic primers used were

His39-F: 5'-A T C T T T G G C A G A C A C T C T G G T C A A-3'

His39-B: 5'-T G A C C A G A G T G T G T C C C A A A G T A-3'

Both genes contained the Cys102Ser mutation as this amino acid replacement has been shown to stabilize *S. c.* cyt *c* and prevent oxidative dimerization.[20,21] To use the PCR method of mutagenizing *cyc1*, two additional primers containing the *Hin*dIII cleavage sequence were required:

*Hin*d-F: 5'-T C C C G G G A G C A A G C T T A A G A T G T T-3'

*Hin*d-B: 5'-T G C A A A T T A A A G C T T T C G A G C G T C-3'

For the His39Ala80Ser102cyt *c* gene target, two PCR reactions were performed, each consisting of 10 μL 2 mM NTP, 1 μL M13mp18 M80A*cyc1*, 10 μM *Hin*d F [*Hin*d B] oligonucleotide (10 pmol/mL), 10 μL His39-B [His39-F] oligonucleotide (10 pmol/μL), 0.5 μl Taq polymerase (5 units/μL), 10 μL 2 mM BMB buffer, and 58.5 μL H_2O to give 100 μL solution. In the second reaction, the components in brackets replaced those proceeding them. Two analogous reactions were performed for the His39Ser102cyt *c* target, with M13mp18*cyc1* substituted for M13mp18 M80A*cyc1*. The reactions were ramped in a thermocycler from 20 to 95 to 20 °C in two minutes, then heated at 95 °C for 20 seconds, 61 °C for 20 seconds and 72 °C for 1 minute. 30 cycles of the 95-61-72 °C sequence were performed, after which the reaction was held at 72 °C for 2 minutes. The PCR products were purified by electrophoresis on a 0.8 % agarose/TAE gel followed by extraction using a QIAEX agarose gel extraction kit from Qiagen. A second PCR reaction for each target was then performed to join these fragments to make the mutant gene. These reactions consisted of 10 μL of each fragment as purified by the QIAEX protocol, 10 μL each of *Hin*d-F and *Hin*d-B oligonucleotides (10 pmol/μL), 10

μM 2 mM NTP, 0.5 μL Taq polymerase (5 units/μL), 10 μM BMB reaction buffer, and 40 μM H_2O. This was run in the thermocycler as described for the above reactions. An analytical 1% agarose TAE gel indicated that ~ 1000 base pair PCR products were obtained as expected for the *cyc1* gene.

The PCR products were cloned using the TA cloning kit for ligation of PCR products from Invitrogen. This method makes use of the overhang A bases typically added onto the 3' end of PCR products and does not require purification of the PCR products prior to ligation. 2 mL of the PCR products were incubated with 2 mL of suspended commercial pCR vector, 41 μL T4 DNA ligase, 1 μL ligation buffer, and 4 μl water at 12 °C for 12 hours. Competent *E. coli* provided with the kit were transformed with the pCR vectors by the heat-shock method as indicated by the manufacturer's instructions. Warm SOC medium was added to the cells and the mixture was shaken at 37 °C for one hour at 225 rpm. The transformed cells were plated onto LB agar plates containing 50 μg/mL ampicillin. Prior to plating, the plates were spread with 25 μL of 40 mg/mL X-Gal in dimethylformamide. The plates were incubated at 37 °C overnight and white colonies were chosen for use in gene expression and sequencing. DNA sequences of the mutated *cyc1* genes were determined by the Sanger dideoxy method.[22]

Gene Expression

The colonies from the TA cloning were grown in LB medium and the plasmids extracted by alkaline lysis of the cells and ethanol precipitation of the DNA. The incorporation of the gene into the plasmid was confirmed with *Hin*dIII cleavage followed by analytical agarose gel electrophoresis. Colonies containing plasmid that had gene incorporation were grown in 500 mL of LB medium. The DNA was extracted and purified by the Qiagen maxi-prep plasmid protocol kit from Qiagen. The pure plasmid was digested with *Hin*dIII and purified by agarose electrophoresis followed by QIAEX extraction of the gene product. The M80A*cyc1* gene was ligated into the YEp213 K

vector (Figure 3.2) which contains a functional (*cyc1+*) gene using T4-DNA ligase. TheSer102*cyc1* gene was ligated into the YEp213 vector. The products were transformed into *E. coli* (strain HB101) by electroporation. A Bio-Rad gene pulser equipped with a Bio-Rad pulse controller was used for all electroporation steps. 40 μL competent *E. coli* were gently mixed with 2.5 mL YEp K and pulsed (25 μF, 2.5 kV, 200 Ω) in an ice-cold, 0.1-cm electroporation cuvette. The cells were immediately suspended in 1 mL SOC medium and transferred to a 17 x 100 mm polypropylene tube and shaken at 37 °C at 225 rpm. After one hour, the cells were plated onto LB agar plates containing ampicillin and incubated at 37 °C. A small selection of the resulting colonies were grown in 500 mL LB medium. Plasmid DNA was prepared from these cells using the Qiagen plasmid kit.

The electroporation method was chosen to transform the yeast cells with the YEp K and YEp plasmids.[23] *S. cerevisiae* GM3C2 (α *leu2-3 leu2-112 trp1-1 his4-519 cyc1-1 cyp3-1*) was grown in YPD medium at 30 °C to an OD_{600} of 1.3. The cells were pelleted by centrifugation (GSA rotor, 5000 rpm, 5 min., 4 °C). The pellet was suspended in 250-mL sterile ice-cold H_2O and the solution was centrifuged (GSA rotor, 5000 rpm, 5 min., 4 °C). This washing and centrifuge step was repeated once more, combining all of the cells into one centrifuge bottle. The cells were resuspended in 20 mL ice-cold 1 M sorbitol and centrifuged (SS-34 rotor, 5000 rpm, 5 min., 4 °C). The supernatant was discarded and the pellet resuspended in 0.2 mL 1 M ice-cold sorbitol. 40 mL of the cell suspension was gently mixed with ~ 500 ng M80AYEp K or YEp and incubated on ice for 5 minutes. This mixture was transferred to a cold 0.2-cm electroporation cuvette and pulsed (25 μF, 1.5 kV, 200 Ω). 1 mL of cold 1 M sorbitol was immediately added to the cuvette and mixed by gentle pipetting. The cell suspension was applied to plates consisting of SD medium with 1 M sorbitol and incubated at 30 °C for 5–7 days. The colonies were replated on YPG to test for iso-1-cyt *c* expression.

Protein Purification

Colonies from the YPG plates were grown in 2 mL of SD medium at 30 °C with rapid (250 rpm) shaking for 3–5 days. The 2-mL preparation was innoculated into 100 mL SD medium. After 5 days of growth at 30 °C with rapid shaking, the 100-mL preparation of media and cells was innoculated into 10 L of YPG medium and grown to saturation (5–6 days). Protein was prepared from 10- or 170-L fermentations of yeast hosts in YPG as described by Smith and coworkers.[21] At each step, the media was supplemented with 50 mg/mL ampicillin to prevent bacterial contamination. Cell lysis was performed by stirring the cells in 1:1 v/w 1 M NaCl, 50 mM NaP_i, 1 mM EDTA and 0.5:1 v/w ethyl acetate for 36 hours at 4 °C. The suspension was centrifuged (GSA rotor, 8000 rpm, 45 min., 4 °C) and the supernatant was brought to 50% saturation with $(NH_4)_2SO_4$ (314 g/L). The supernatant was centrifuged (GSA rotor, 8000 rpm, 45 min., 4 °C), filtered through a course glass frit, and dialyzed at least three times against 10 volumes of 12.5 mM NaP_i, 1 mM EDTA (pH 7.0). The contents of the dialysis bags were added to 300 mL (or 500 mL for purification of protein from large-scale fermentations) of fast-flow CM52 Sepharose ion-exchange resin prewashed with an equal volume of 50 mM NaP_i, pH 7.0. The resin was poured into a 5 x 20 cm Econo-column (Bio-Rad) and washed with 1 column volume of 50 mM NaP_i, pH 7.0. The protein was eluted with ~ 2 column volumes of a 0–1 M NaCl gradient in 50 mM NaP_i (pH 7.0). The nonfunctional Ala80 gene product was separated from the functional (*cyc*+) product which contains the di-His chelating site (Figure 3.3) by immobilized metal affinity chromatography. A 5 mL Hi-Trap iminodiacetic acid (IDA) column (Pharmacia) was charged with 2.5 mL of 0.1 M $CuSO_4$ according to the manufacturer's instructions. The cytochromes were loaded onto the column in a volume < 10 mL in 50 mM NaP_i, 500 mM NaCl, 1 mM imidazole (pH 7.0) and the nonfunctional product eluted with the void column volume. A gradient of 1–20 mM imidazole was run to elute the functional

chelating cytochrome (Figure 3.4). The proteins were oxidized overnight with K[Co(EDTA)] and purified by FPLC on a Mono S HR 10/10 column. The product was eluted by running a 0–0.2 M NaCl gradient (70 mL) followed by a 0.2–0.4 M NaCl gradient (170 mL). The mutant cytochrome eluted at ~ 0.3 M NaCl (Figure 3.5).

Characterization of Cytochrome *c* Variants

The introduction of the Met80Ala mutation into Ala80cyt *c* was verified by peptide sequencing. The holoprotein (0.7 mM solution in 70% formic acid) was cleaved with CNBr (130x excess) for 46 hours. The reaction was terminated by dilution with 500x excess H_2O followed by repeated ultrafiltration into 1000x excess $AgNO_3$ and dilution with 100 mM NH_4HCO_3 (pH 8.0) until the pH of the filtrate was 8.0. The sample was lyophilized, dissolved in H_2O, and chromatographed on a Pharmacia ProRPC HR 5/10 column to isolate the 65–103 fragment. Sites of heme attachment (Cys 14 and 17) and the mutation of Met80 to Ala were confirmed by N-terminal sequencing of the holoprotein and of the 65–103 CNBr fragment. Sequencing was done by automated Edman degradation on an Applied Biosystems 473A sequencer.

The procedures and instrumentation for SDS/PAGE electrophoresis, absorption spectroscopy, CD spectroscopy, MCD spectroscopy, preparation of axially-ligated derivatives, and measurement of the O_2 affinity and the autoxidation rate were described in Chapter 2.

Molecular Modeling

Energy-minimization calculations on His39/Leu58His cyt *c* were performed on the 1.23-Å structure of yeast iso-1-cytochrome c[24] using BIOGRAF version 3.0 (Molecular Simulations).

Figure 3.4 Chromatogram showing the purification on a Cu(II)IDA column of homogenized protein extracts containing the nonfunctional *S. c.* Ala80cyt *c* and the functional *S. c.* His39/Leu58Hiscyt *c* proteins. About 50 mg of protein extracts were loaded onto a 5-mL Hi-Trap Chelating column (Pharmacia LKB) that was charged with 2 mL of 100 mM $CuSO_4$. Ala80cyt *c* eluted with the void volume and the chelating cytochrome was eluted with a 1–20 mM imidazole gradient in 50 mM NaP_i, 500 mM NaCl (pH 7.0). Peak I contains the Met80Ala mutant and other proteins that bind weakly to the column. Peaks II and III contain the reduced and oxidized *S. c.* His39/Leu58His proteins, respectively.

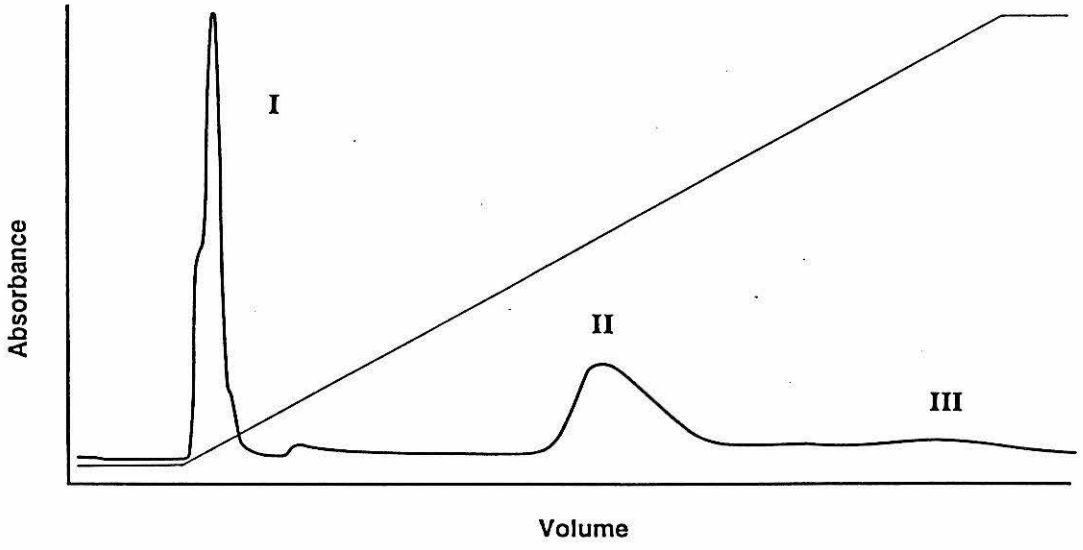
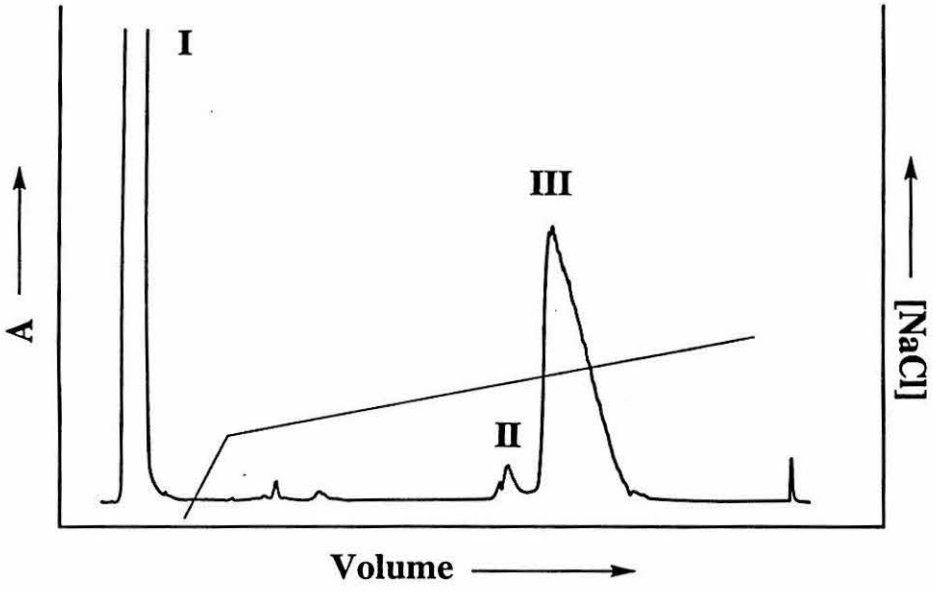

I
II
III
Absorbance
Volume

Figure 3.5 Typical cation-exchange (FPLC Mono S HR 10/10) chromatogram for the purification of *S. c.* Ala80cyt *c*. The sample was oxidized with K[Co(EDTA)] prior to purification. The K[Co(EDTA)] elutes with the void volume (peak I). Peak III is pure *S. c.* Fe(III)Ala80cyt *c*, preceded by a minor component (peak II) believed to be deamidated cyts *c*. The thin line represents the gradient of NaCl, the shallow part of which runs from 200 to 400 mM NaCl. Detection is at 280 nm on the 1.0 absorbance scale.


I
III
II
A
[NaCl]
Volume

Results and Discussion

Characterization of Functional *S. c.* Cytochromes *c*

Functional cyts *c* refers to the cytochromes *c* expressed that have the native Met at position 80. One of these mutants is that which is co-expressed with the nonfunctional Ala80cyt *c*, His58Ser102cyt *c*, which has the di-histidine (His58 and the native His39) chelating site on the surface. The characterization of this product is described elsewhere.[17,19] Because His58Ser102cyt *c* exhibits a lower thermal stability and higher reduction potential than native protein, Ser102cyt *c* was expressed for the purpose of providing "wild-type" protein to which to compare the Ala80cyt *c* variants. Typically, 2–4 mg/mL of Ser102cyt *c* were obtained from 10 L preparations of *S. c.* cells, in accordance with yields of cytochrome reported elsewhere.[13,21,25] The characterization of Ser102cyt *c* is described elsewhere.[20] The protein expressed here was found to exhibit chromatographic and electrophoretic properties consistent with its 12-kDa molecular weight. In addition, the UV/vis and CD spectra were identical with those reported for both Ser102 and Thr102cyt *c*.

Characterization of *S. c.* Gln39Ala80Ser102cyt *c*

Because the focus of this work is on the effect of the mutation of the Met80 ligand to Ala, for brevity, S. *c.* Gln39Ala80Ser102cyt *c* will be referred to as *S. c.* Ala80cyt *c*, or as Ala80cyt *c* when it is clear that it is of the *S. c.* species. His39 is a surface residue and it has already been shown that its mutation to Gln has no affect on the properties (spectroscopy, reduction potential, fold) of *S. c.* iso-1-cyt *c*.[20] The native Cys102 residue was replaced by Ser because it has been shown that this stabilizes the protein, preventing oxidative dimerization, and has little affect on the protein structure.[20] As Ser102cyt *c* (or Thr102cyt *c*) is used almost universally in studies of *S. c.* cyt *c*, Ser102cyt *c* will be referred to as wild-type (WT) cyt *c*. Although Ala80cyt *c* has been expressed with the native His39 in place, this work focused on the characterization of the Gln39Ala80Ser102

mutant because the levels of expression were higher than for His39Ala80Ser102. In all preparations of nonfunctional cytochromes, ~ 2 mg of cytochromes were obtained per liter, a level of expression comparable to that of other expression systems used to generate functional iso-1-cyt *c* variants.[13,21,25] The Gln39Ala80Ser102 protein constitutes ~ 25% of the total cytochromes *c* to give ~ 0.5 mg/liter expression. In the case of the Ala80 variant with the native His39 in place, less than 20% of the cytochromes expressed were the nonfunctional variant.

Chromatography and Electrophoresis

The FPLC cation exchange chromatogram of *S. c.* Ala80cyt *c* is shown in Figure 3.5. This purification procedure easily separates the product from what is believed to be deamidated forms of the protein which elute before the major peak. In addition, it is possible to chromatographically separate oxidized and reduced forms of the protein, although best results are obtained when the sample is fully oxidized with K[Co(EDTA)] before purification. The fractions taken during the elution of the product have different spectroscopic characteristics from each other. The fractions eluting early have UV/vis spectra characteristic of the neutral pH form of the oxidized protein (*vide infra*) while the later fractions have spectra similar to the low pH form of the protein. It was confirmed that this difference is not due to a difference in pH among the fractions, but rather to the change in ionic strength during the elution. Exchanging the buffer (to either 25 or 50 mM NaP_i (pH 7)) by gel filtration gives a homogeneous sample with spectral characteristics identical to those of the neutral pH protein, as described below.

The molecular weight of *S. c.* Ala80 was found to be the same as that of wild-type by SDS/PAGE analysis. CZE analysis confirmed that a homogeneous cyt *c*-like product was isolated.

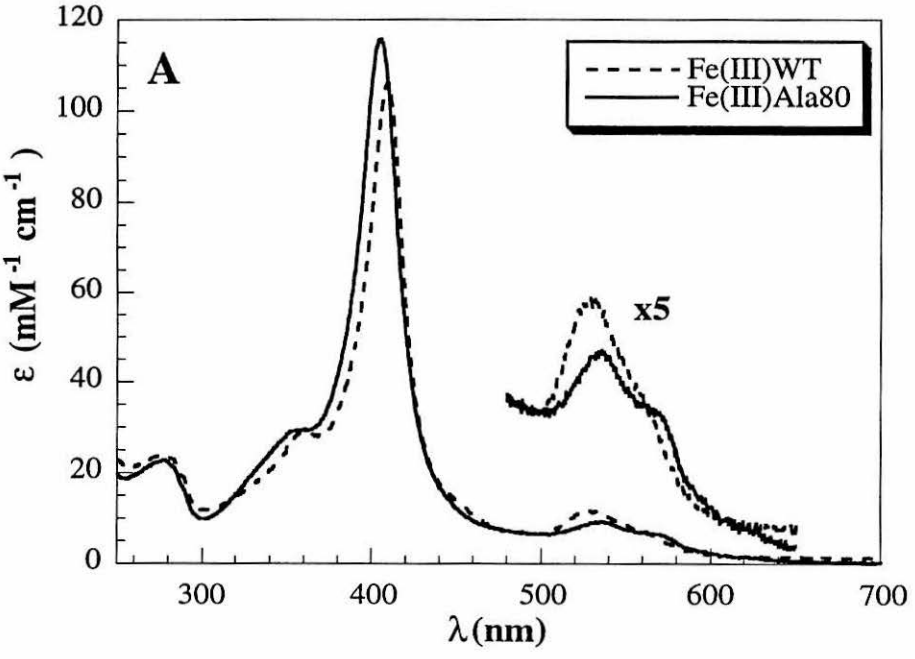
Absorption Spectroscopy

The absorption spectra of oxidized and reduced *S. c.* Ala80 under aerobic conditions are markedly different from those of wild-type (Figure 3.6) and similar to those of semisynthetic Ala80cyt *c*. As is seen for the semisynthetic protein, the Soret band of Fe(II)Ala80 is blue-shifted from that of reduced wild-type and the α-band intensity is significantly reduced. Fe(III)Ala80 has a blue-shifted Soret band compared to wild-type and, like semisynthetic Ala80, three poorly-defined bands at lower energy (536, 568, and 626 nm). As expected, the 695 nm Met80 → Fe(III) charge-transfer band is absent in the spectrum of the mutant. The differences from wild-type confirm that a change in heme coordination has occurred and the similarities to semisynthetic Ala80 strongly suggest that, at pH 7.0 under aerobic conditions, *S. c.* Fe(III)Ala80 has His-OH^- ligation and *S. c.* Fe(II)Ala80 has dioxygen bound to the heme.

Circular Dichroism Spectroscopy

The far-UV CD spectrum of *S. c.* Fe(III)Ala80cyt *c* is similar to that of *S. c.* WT Fe(II)cyt *c* (Figure 3.7), indicating that the α-helical content of *S. c.* Ala80 is nearly the same as that of WT. The thermal stability of the mutant was determined by monitoring the change in CD at 222 nm with temperature (Figure 3.8). Wild-type cyt *c* has an unfolding transition temperature of 55 °C. As is evident in Figure 3.8, the CD melting curve of Fe(III)Ala80 at pH 7 shows 2 transitions at 35 and 55 °C. Lowering the pH to 5.5 gives one transition at 50 °C, and raising the pH to 8 gives one transition at 58 °C. It is possible that deprotonation of a heme-bound water affects the temperature dependence of the fold of *S. c.* Fe(III)Ala80; at pH 8, hydroxide is bound, giving a neutral heme core and stabilizing the hydrophobic core of the protein. At lower pH, however, protonation of the hydroxide results in the addition of an uncompensated charge in the hydrophobic core. The two transitions seen at pH 7 could be the results of a change in a temperature dependence of the pK_a of the heme-bound water or other ionizable groups that affect the

Figure 3.6 UV/vis absorption spectra of *S. c.* Ala80cyt *c* (solid lines) and *S. c.* Ser102 cyt *c* (dashed lines) in the (A) oxidized and (B) reduced states. The reduced form of Ala80cyt *c* has dioxygen bound to the heme, and the oxidized form has been identified to have hydroxide ligation at the heme. Spectra are recorded on 5–7 μM protein samples in 25 mM NaP_i (pH 7.0) in a 1-cm cell.


A
Fe(III)WT
Fe(III)Ala80
x5
ε (mM-1 cm-1)
λ(nm)
0
20
40
60
80
100
120
300
400
500
600
700

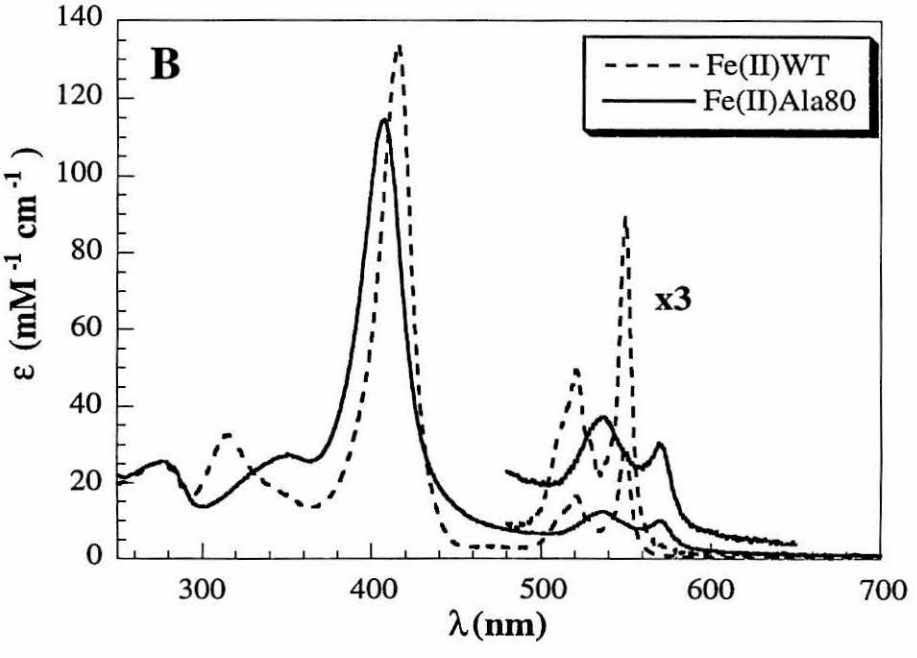

B
Fe(II)WT
Fe(II)Ala80
x3
ε (mM-1 cm-1)
λ(nm)
0
20
40
60
80
100
120
140
300
400
500
600
700

Figure 3.7 Far-UV CD spectra of *S. c.* Fe(III)WTcyt *c* at pH 7.0 (dotted line) and *S. c.* Fe(III)Ala80cyt *c* at pH 7.0 (solid line) and at pH 8.0 (dashed line). Samples were 12 μM in 25 mM NaP_i (pH 7.0) in a 0.1-cm cell.

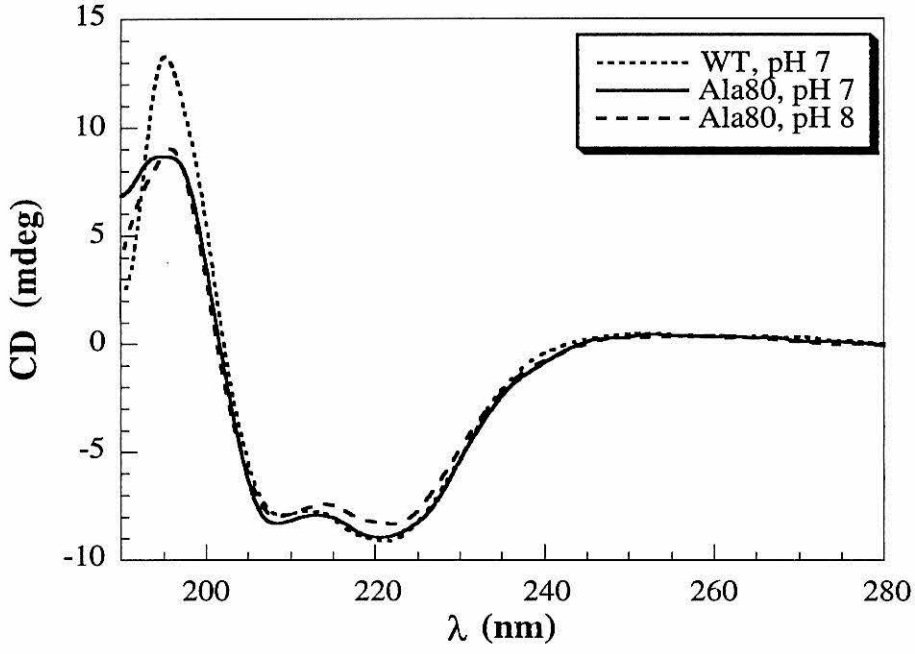
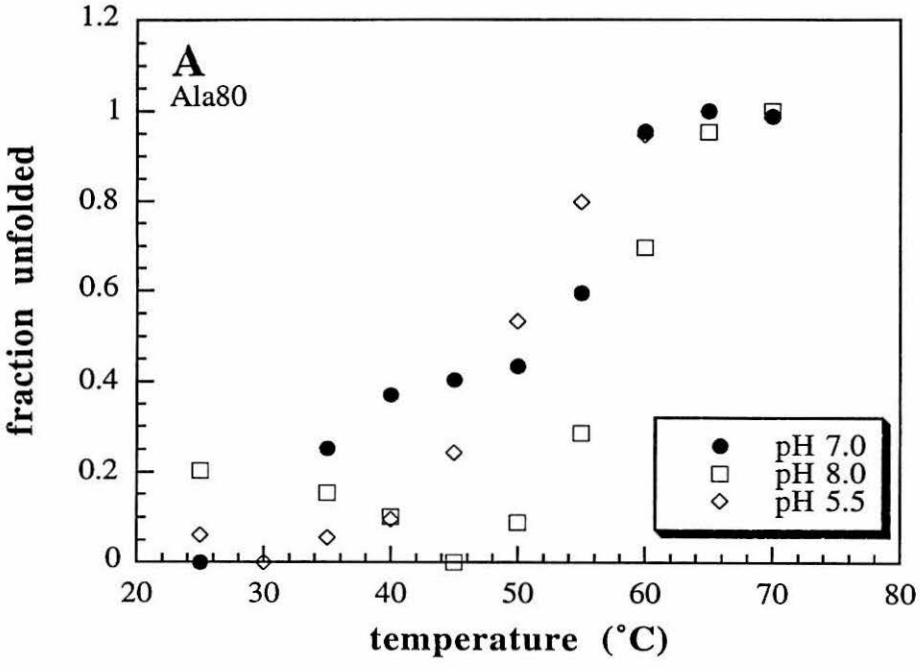
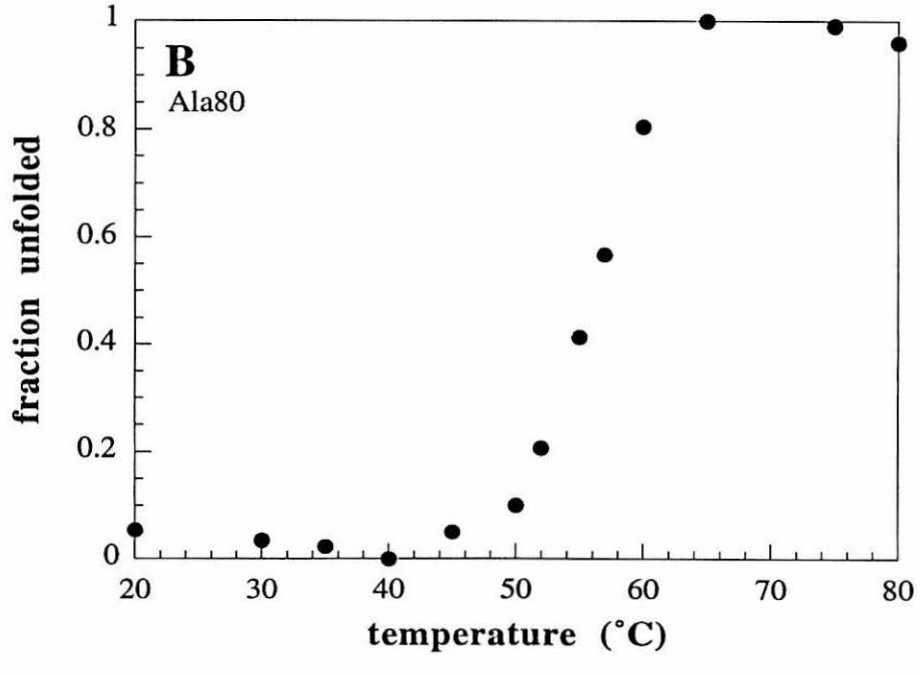

15
10
5
0
-5
-10
CD (mdeg)
200
220
240
260
280
λ (nm)
WT, pH 7
Ala80, pH 7
Ala80, pH 8

Figure 3.8 CD melting curves for (A) *S. c.* Fe(III)Ala80cyt *c* at pH 5.5 (diamonds), pH 7.0 (circles), and pH 8.0 (squares) and (B) *S. c.* Fe(III)WTcyt *c* (pH 7.0). The protein samples were 12 μM in 25 mM NaP_i. The pH of the Ala80 sample was adjusted by addition of small amounts of 30% H_3PO_4 or 2 M NaOH.


A
Ala80
fraction unfolded
temperature (°C)
pH 7.0
pH 8.0
pH 5.5
0
0.2
0.4
0.6
0.8
1
1.2
20
30
40
50
60
70
80


B
Ala80
fraction unfolded
temperature (°C)
0
0.2
0.4
0.6
0.8
1
20
30
40
50
60
70
80

protein fold. It is well-known that the stability of cyt *c* is highly dependent on is oxidation state, with the reduced protein showing enhanced stability.[26] Here, a change of heme ligation and, thus, net charge at the heme site, may give an analogous effect.

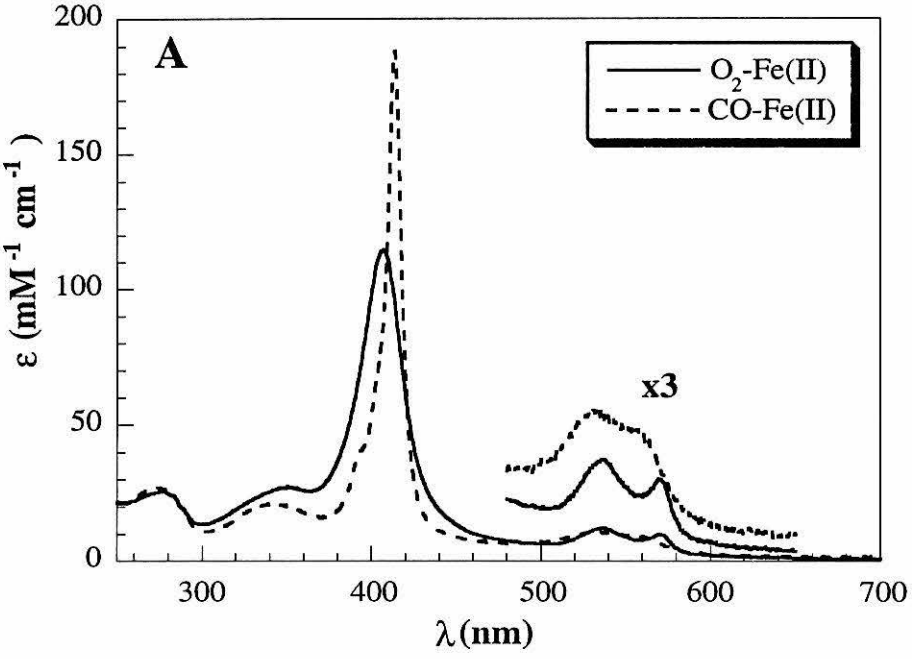
Ligand-Binding to *S. c.* Gln39Ala80Ser102cytochrome *c*

Absorption Spectroscopy

The UV/vis absorption spectra of *S. c.* Ala80 are given in Figures 3.9 (Fe(II) derivatives) and 3.10 (Fe(III) derivatives). Spectral maxima and extinction coefficients are presented in Table 3.1. The CO and O_2 derivatives (Figure 3.9A) have spectral features very similar to those of both Mb and of semisynthetic Ala80cyt *c*. *S. c.* deoxy-Fe(II)Ala80 (Figure 3.9B), however, has a markedly different spectrum from that of horse heart deoxy-Fe(II)Ala80. The low-energy shoulder on the Soret band of *S. c.* deoxy-Fe(II)Ala80 has a lower intensity than that seen for horse heart deoxy-Fe(II)Ala80. The α, β region is also quite different, with *S. c.* deoxy-Fe(II)Ala80 exhibiting a relatively intense ($\varepsilon = 14.8\ mM^{-1}cm^{-1}$) band at 548 nm. As with the semisynthetic protein, the spectrum of *S. c.* deoxy-Fe(II)Ala80 appears to have two (or more) components that may be due to high- and low-spin heme. In the yeast protein, however, the low-spin component appears to be the major one.

The UV/vis spectrum of the *S. c.* Ala80 in the oxidized state exhibits a strong pH dependence, giving spectra characteristic of a low-spin heme above pH 7, and of a high-spin heme below pH 7 (Figure 3.10A). The spectral changes are similar to those observed for semisynthetic Fe(III)Ala80, suggesting that at and above neutral pH, *S. c.* Fe(III)Ala80 has hydroxide bound to the heme, and that this hydroxide ligand is protonated at acidic pH values. The extinction coefficient ($\varepsilon = 142.9\ mM^{-1}cm^{-1}$) of the Fe(III)Ala80 Soret band at low pH is characteristic of a high-spin hexacoordinate ferric heme.[27]

Figure 3.9 (A) UV/vis absorption spectra of the (A) O_2 (solid line) and CO (dashed line) adducts of *S. c.* Fe(II)Ala80cyt *c*. (B) UV/vis absorption spectrum of the deoxygenated form of *S. c.* Fe(II)Ala80cyt *c*. Note the differences between this spectrum and that of the ferrous deoxy form of semisynthetic Ala80cyt *c* in Figure 2.20A. The samples were 8 μM in protein in 25 mM NaP_i (pH 7.0) in a 1-cm cell.


A
O2-Fe(II)
CO-Fe(II)
x3
ε (mM-1 cm-1)
λ (nm)
0
50
100
150
200
300
400
500
600
700

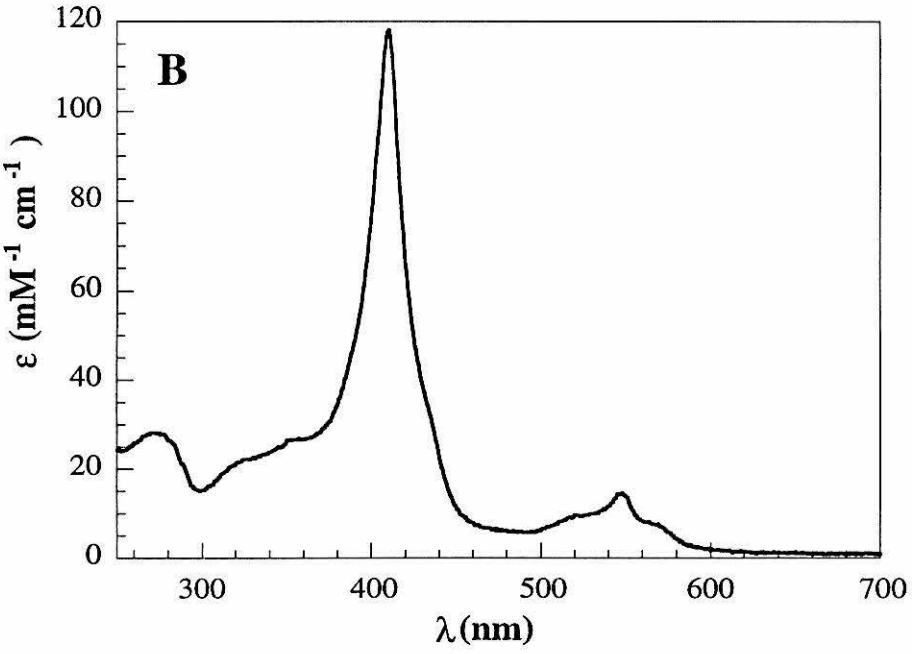
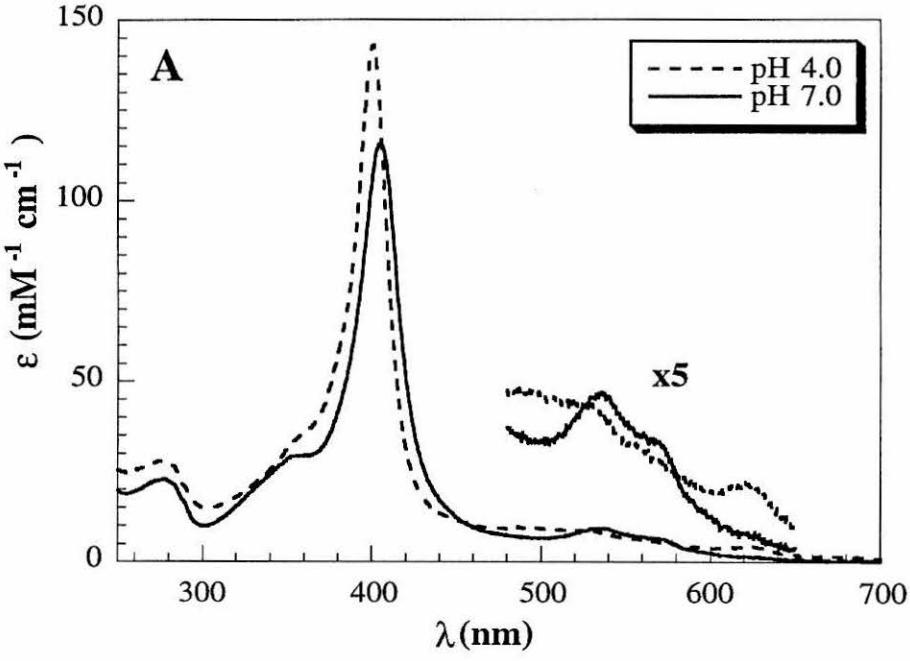

B
ε (mM-1 cm-1)
λ (nm)
0
20
40
60
80
100
120
300
400
500
600
700

Figure 3.10 (A) UV/vis absorption spectra of *S. c.* Fe(III)Ala80cyt *c* at pH 7.0 (solid line) and 4.0 (dashed line). (B) UV/vis absorption spectrum of the cyanide adduct of *S. c.* Fe(III)Ala80cyt *c* at pH 7.0. Spectra were obtained on 6–8 μM in protein in 25 mM NaP_i (pH 7.0), or 25 mM NaOAc (pH 4.0) in a 1-cm cell.


A
pH 4.0
pH 7.0
x5
150
100
50
0
ε (mM-1 cm-1)
300
400
500
600
700
λ(nm)

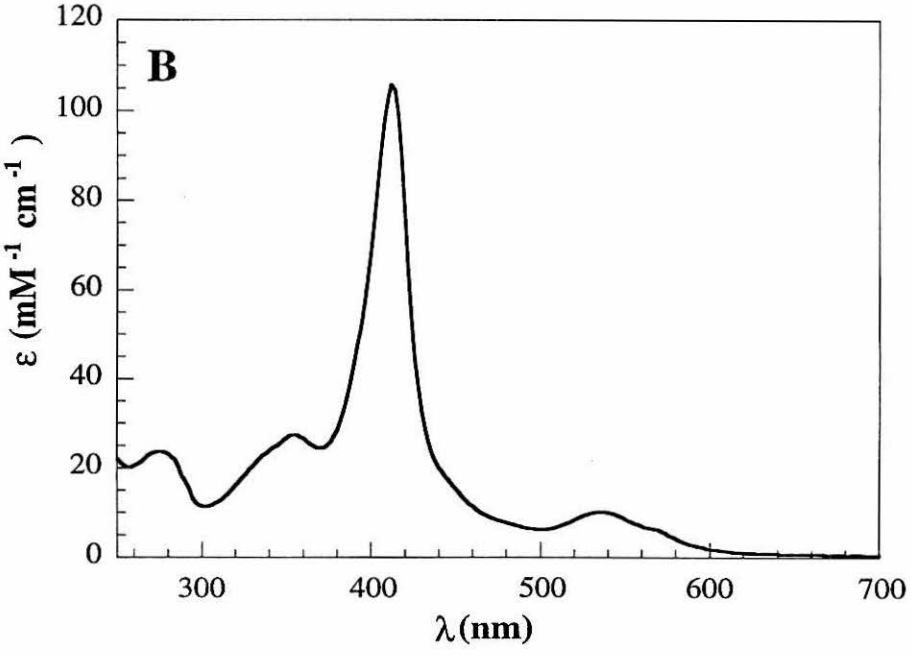

B
120
100
80
60
40
20
0
ε (mM-1 cm-1)
300
400
500
600
700
λ(nm)

Table 3.1

UV/vis absorption data for derivatives of S. c. Ala80cyt c[a]

	α		β	γ (Soret)	δ		UV
ferrous derivatives:							
deoxy-Ala80	564[b]	548	521[b]	410/437[b]	357[b]	327	270
	8.3	*14.8*	*9.9*	*118.2*[c]	*29.5*	*22.4*	*28.4*
O_2-Ala80	570		537	408	351		275
	10.1		*12.5*	*114.8*	*27.7*		*25.8*
CO-Ala80	554		531	414	345		275
	9.9		*11.1*	*188.2*	*21.3*		*27.1*
ferric derivatives:							
H_2O-Ala80 (pH 4.0)	626		537	401	352		275
	4.2		*8.0*	*142.9*	*32.4*		*28.0*
HO-Ala80 (pH 7.0)	626	568	536	406	357		278
	1.6	*6.9*	*9.4*	*115.8*	*29.5*		*22.9*
CN-Ala80		536		412	354		274
		10.3		*105.8*	*27.5*		*23.8*

[a]Upright figures indicate wavelength in nm, italics indicate millimolar extinction coefficient.

[b] Shoulders.

[c]ε at 410 nm.

Magnetic Circular Dichroism Spectroscopy

The MCD spectrum of the low-pH (pH 4.5) form of *S. c.* Fe(III)Ala80 (Figure 3.11B) is similar to those of semisynthetic H_2O-Fe(III)Ala80 and of Fe(III)Mb[28] (Figure 2.27), confirming the assignment of His/H_2O-ligation to a high-spin heme made according to absorption data. The spectrum of *S. c.* Fe(III)Ala80 at pH 7 (Figure 3.11A), however, also appears to be that of a high-spin heme rather than low-spin as was expected based on the analogy to h. h. Ala80 and on the UV/vis absorption spectrum. The reason

Figure 3.11 MCD spectra of *S. c.* Fe(III)Ala80cyt *c* at 4 °C. (A) 11 μM sample in a 1-cm pathlength microcell at pH 7. (B) 31 μM sample in a 0.2-cm pathlength cell at pH 7. The spectra indicate that the protein is high-spin at both of these pH values.

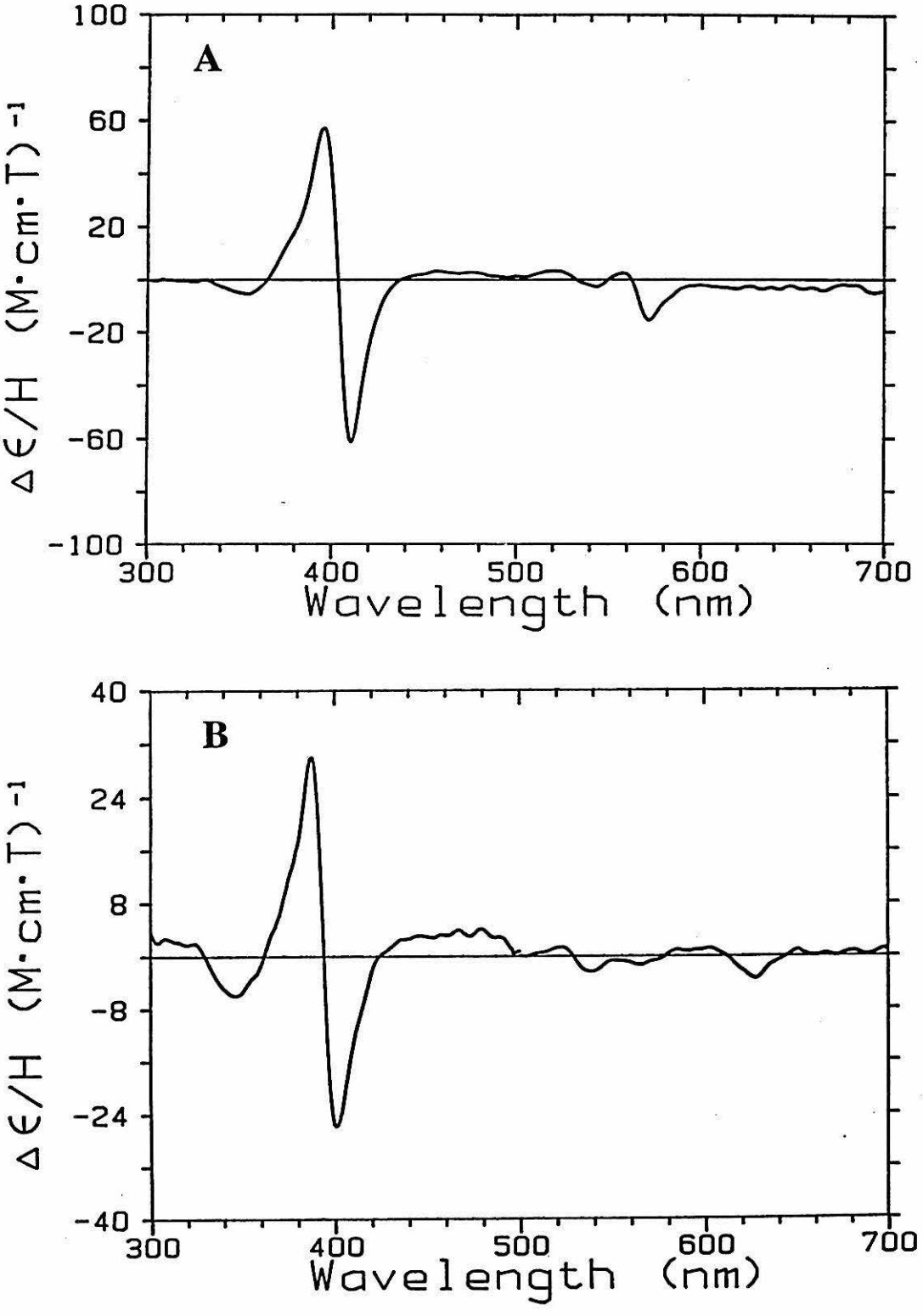
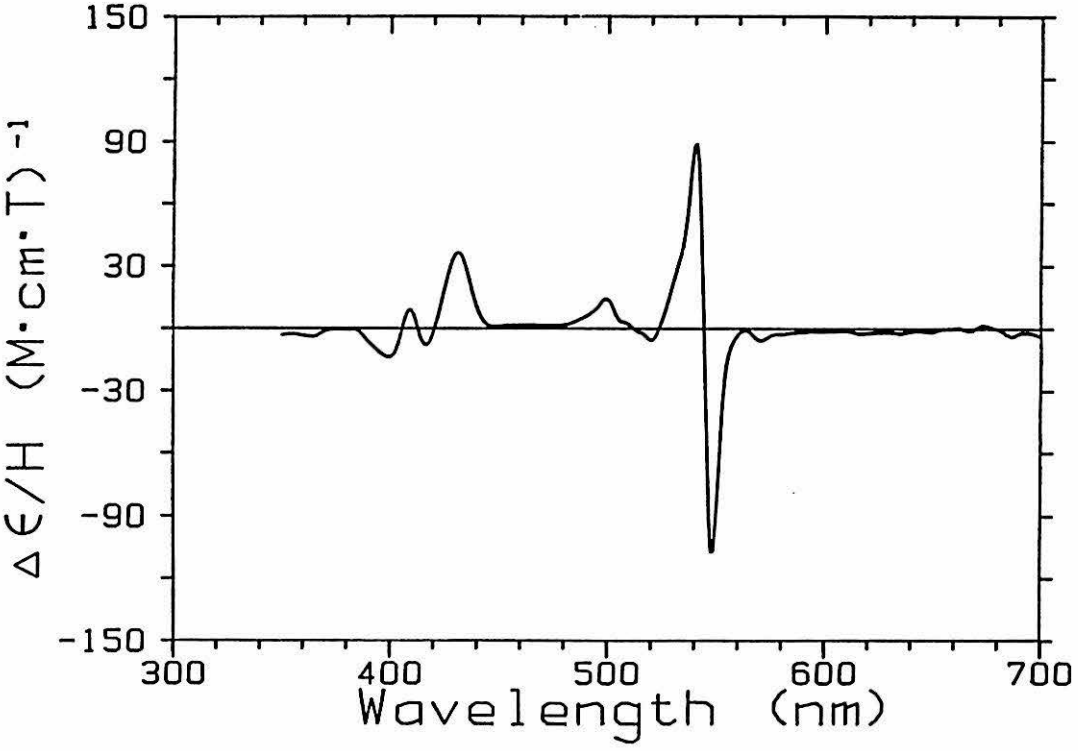

A
100
60
20
-20
-60
-100
300
400
500
600
700
Wavelength (nm)
ΔE/H (M·cm·T)⁻¹
B
40
24
8
-8
-24
-40
300
400
500
600
700
Wavelength (nm)
ΔE/H (M·cm·T)⁻¹

Figure 3.12 MCD spectrum of *S. c.* deoxy-Fe(II)Ala80cyt *c* (11 μM in a 1-cm pathlength microcell) in 25 mM NaP_i (pH 7.0) at 4 °C. The spectrum is typical for that of a ferrous heme in the reduced state, with a small high-spin component.


150
90
30
-30
-90
-150
300
400
500
600
700
Δε/H (M·cm·T)⁻¹
Wavelength (nm)

for this apparent discrepancy between the MCD and absorption data is unclear; however, MCD is a more definitive method for determining spin-state.

S. c. deoxy-Fe(II)Ala80 has an MCD spectrum with a feature in the α, β region indicative of high-spin ferrous heme (Figure 3.12), as seen by comparison to the spectrum of deoxy-Mb (Figure 2.26B).[28] In addition, a small band is found in the Soret region which suggests the presence of high-spin Fe(II) heme. The MCD data therefore confirm that *S. c.* deoxy-Fe(II)Ala80 is mostly low-spin.

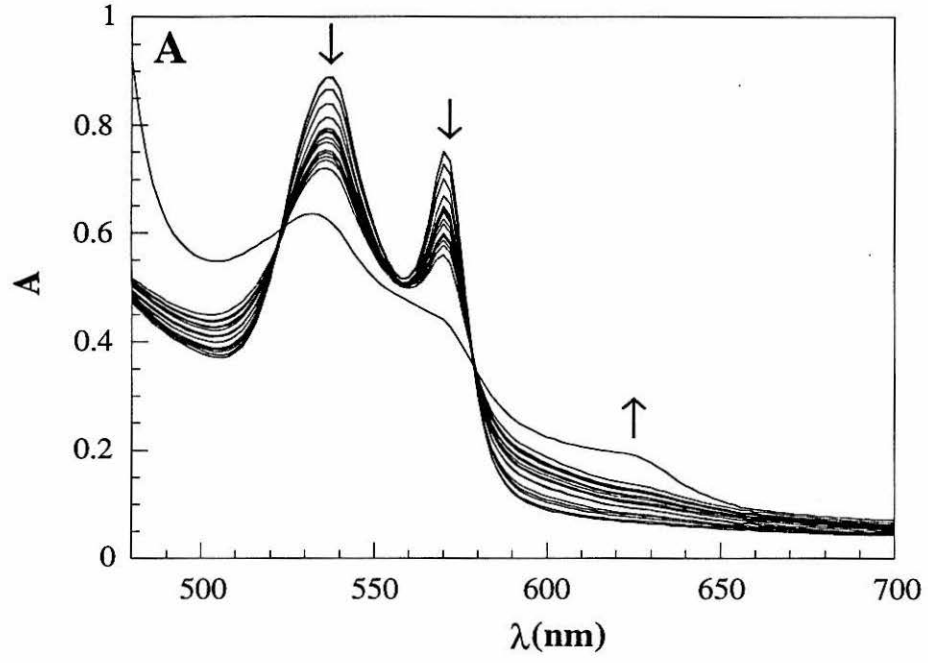
Ligand Photolysis

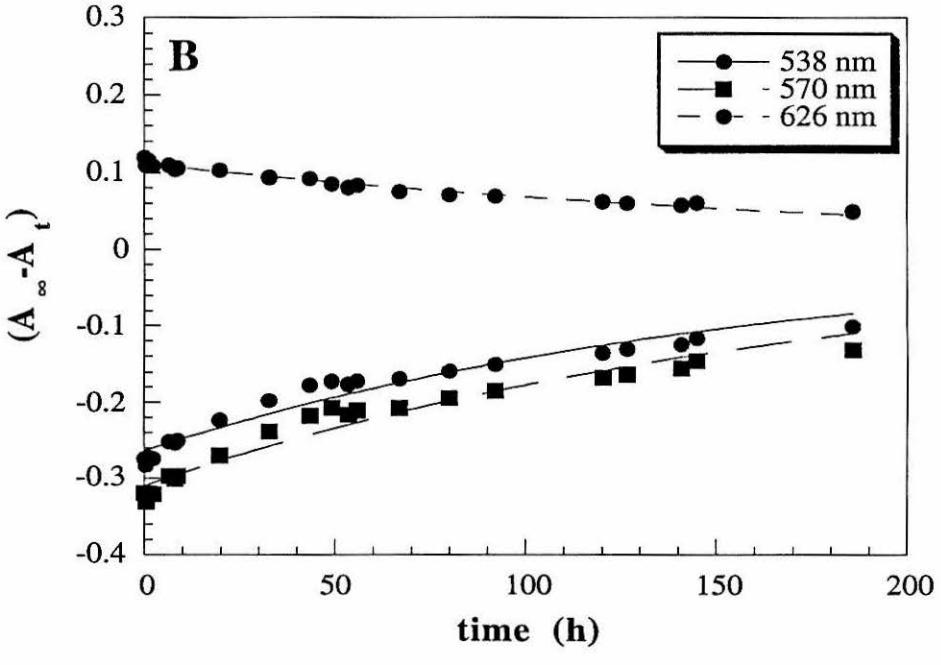
A number of attempts to observe the binding of CO after photolysis from *S. c.* Fe(II)Ala80 were made; however, no signal was observed monitoring either at the Soret band for *S. c.* deoxy-Fe(II)Ala80 or for CO-Fe(II)Ala80. Therefore, no photodissociation of CO from *S. c.* Fe(II)Ala80 occurs on timescales observable using the nanosecond transient absorption system, indicating a lower quantum yield for this process than was detected for semisynthetic Ala80. The signal size observed in the transient absorption spectrum for the semisynthetic mutant was quite small and near the threshold of what can be detected using the nanosecond transient absorption system.

Autoxidation

The dioxygen adduct of *S. c.* Fe(II)Ala80 was found to be remarkably resistant to autoxidation. The autoxidation rate determined by monitoring the absorption spectrum (Figure 3.13) was found to be 0.0030(7) h^{-1}, giving it a half-life of 9.6 days at 22 °C. This autoxidation rate is among the lowest reported for heme proteins and is on the order of that of Leu29Phe sperm whale Mb (k = 0.001 h^{-1} at 37 °C). The greater resistance of *S. c.* O_2-Fe(II)Ala80cyt *c* to autoxidation relative to semisynthetic O_2-Fe(II)Ala80 (k = 0.010(3) h^{-1} at 22 °C; see Table 2.5) may be related to the decreased quantum yield of CO dissociation from the yeast protein relative to the horse heart protein; the yeast cyt *c* heme

Figure 3.13 (A) UV/vis absorption spectra taken over the course of the autoxidation of *S. c.* O_2-Fe(II)Ala80cyt *c*. Data were recorded on 80 μM protein in 25 mM NaP_i (pH 7.0) at 22 °C. The arrows indicate the direction of spectral changes as the sample oxidizes. (B) Plot of autoxidation data. The quantity $(A_\infty - A_t)$ is plotted against time and the data fit to a single exponential to obtain the autoxidation rate, 0.0030(7) h^{-1}.


A
A
λ(nm)
1
0.8
0.6
0.4
0.2
0
500
550
600
650
700


B
538 nm
570 nm
626 nm
(A∞-At)
time (h)
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
0
50
100
150
200

is even less accessible to small molecules than that of horse heart cyt *c*. In addition, the nature of dioxygen-heme pocket interactions may be different between the two proteins; however, most of the heme pocket residues, in particular Phe82 and Tyr67, are conserved between the two species and this is unlikely. Subtle differences in the heme pocket structure between the two proteins may account for some of the enhanced stability of *S. c.* O_2-Fe(II)Ala80. Unlike the data for semisynthetic O_2-Fe(II)Ala80, the UV/vis spectra during the autoxidation of *S. c.* O_2-Fe(II)Ala80 maintain isosbestic points throughout the experiment, indicating a simple two-state transition.

Dioxygen Affinity

The affinity of Fe(II)Ala80cyt *c* for dioxygen was found to be similar to that observed for the semisynthetic Fe(II)Ala80, with an association constant $K = 3.0(1.0)$ μM^{-1}. The large uncertainty is a result of the long time required to reach equilibrium after each addition of dioxygen, leading to some sample autoxidation during the course of the experiment. Due to the difficulty of determining dioxygen affinity for Met80Ala cytochromes *c*, it would be misleading to make explicit comparisons to various dioxygen binders; however, it is clear that *S. c.* Fe(II)Ala80 has a dioxygen affinity on the order of that seen for dioxygen-storage and transport heme proteins in nature (Table 2.6).

Conclusions

Site-directed mutagenesis has been shown to be a viable method for preparing multi-milligram quantities of *S. c.* Ala80cyt *c*. Gel electrophoresis, column chromatography, CZE, and amino acid sequencing indicate that a cyt *c*-like product is formed with the appropriate heme-polypeptide linkages in place. CD spectroscopy confirms that the secondary structure of *S. c.* Ala80cyt *c* is similar to that of the native protein. The thermal unfolding of *S. c.* Fe(III)Ala80 was found to exhibit a dependence on pH, with the protein at pH 8.0 giving a melting curve similar to that of *S. c.* WTcyt *c*,

but at pH 7 exhibiting two transitions. At pH 5.5, the mutant has a lower thermal stability than wild-type. This behavior is attributed to the affect of the protonation state of the heme-bound H_2O/OH^-.

S. c. Ala80cyt *c* binds small ligands to give UV/vis absorption spectra in analogy to those of Mb and of h. h. Ala80cyt *c*. While ferrous-deoxy Mb is high-spin, however, the UV/vis absorption and MCD spectra of *S. c.* deoxy-Fe(II)Ala80 indicate a mostly low-spin heme. Although the *S. c.* Fe(III)Ala80 UV/vis spectra have a pH dependence similar to that observed for the h. h. Fe(III)Ala80, MCD spectroscopy indicates that, even at pH 7.0, oxidized *S. c.* Ala80 is high-spin. This is in contrast to the MCD data for h. h. Ala80 which indicate a low-spin heme under the same conditions.

As was observed for h. h. Fe(II)Ala80, *S. c.* Fe(II)Ala80 binds dioxygen with an affinity similar to or greater than that reported for sperm whale Mb and other mammalian Mbs. The autoxidation rate of the yeast variant, however, is even slower than that of horse heart Ala80. The slower autoxidation rate of *S. c.* O_2-Fe(II)Ala80 is attributed to greater protection of the heme from solvent as indicated by the inability to observe CO photolysis from CO-Fe(II)Ala80.

References and Notes

(1) The work in this chapter has been done in collaboration with Prof. Yi Lu, Dr. Danilo R. Casimiro, and Prof. Jack H. Richards.

(2) Ferrer, J. C.; Turano, P.; Banci, L.; Bertini, I.; Morris, E. K.; Smith, K. M.; Smith, M.; Mauk, A. G. *Biochemistry* **1994,** *33,* 7819-7829. McRee, D. E.; Jensen, G. M.; Fitzgerald, M. M.; Siegel, H. A.; Goodin, D. B. *Proc. Natl. Acad. Sci. U.S.A.* **1994,** *91,* 12847-12851. Goodin, D. B.; McRee, D. E. *Biochemistry* **1993,** *32,* 3313-3324. Vitello, L. B.; Erman, J. E.; Miller, M. A.; Mauro, J. M.; Kraut, J. *Biochemistry* **1992,** *31,* 11524-11535. Satterlee, J. D.; Erman, J. E.; Mauro, J. M.; Kraut, J. *Biochemistry* **1990,** *29,* 8797-8804.

(3) Mizoguchi, T. J.; DiBilio, A. J.; Gray, H. B.; Richards, J. H. *J. Am. Chem. Soc.* **1992,** *114,* 10076-10078.

(4) Mizoguchi, T. J., Ph.D. Thesis, California Institute of Technology, submitted December, 1995.

(5) Raphael, A. L.; Gray, H. B. *Proteins* **1989,** *111,* 766-767. Raphael, A. L.; Gray, H. B. *J. Am. Chem. Soc.* **1991,** *113,* 1038-1040.

(6) Wallace, C. J. A.; Clark-Lewis, I. *J. Biol. Chem.* **1992,** *267,* 3852-3861.

(7) Bren, K. L.; Gray, H. B. *J. Am. Chem. Soc.* **1993,** *115,* 10382-10383.

(8) Wallace, C. J. A. *Curr. Opin. Biotech.* **1995,** *6,* 403-410. Humphries, J.; Offord, R. E.; Smith, R. A. G. *Curr. Opin. Biotech.* **1991,** *2,* 539-543. Offord, R. E. *Prot. Eng.* **1987,** *1,* 151-157.

(9) Wallace, C. J. A. *FASEB J.* **1993,** *7,* 505-515.

(10) Vita, C.; Gozzini, L.; Di Bello, C. *Eur. J. Biochem.* **1992,** *204,* 631-640. Di Bello, C.; Vita, C.; Gozzini, L. *Biochem. Biophys. Res. Commun.* **1992,** *183,* 258-264.

(11) Cafferty, M. S.; Cusanovich, M. A. *Biochim. Biophys. Acta* **1994,** *1187,* 277-288.

(12) Garfinkel, D. J.; Strathern, J. N. *Methods Enzymol.* **1991,** *194,* 342-361.

(13) Pielak, G. J.; Mauk, A. G.; Smith, M. *Nature* **1985,** *313,* 152-154.

(14) Hampsey, D. M.; Das, G.; Sherman, F. *FEBS Lett.* **1988,** *231,* 275-283.

(15) Garcia, L. L.; Fredericks, Z.; Sorrell, T. M.; Pielak, G. J. *New J. Chem.* **1992,** *16,* 629-632. Sorrell, T. N.; Martin, P. K.; Bowden, E. F. *J. Am. Chem. Soc.* **1989,** *111,* 766-767.

(16) Lu, Y.; Casimiro, D. R.; Bren, K. L.; Richards, J. H.; Gray, H. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993,** *90,* 11456-11459.

(17) Umana, P.; Kellis, J. T.; Arnold, F. H. In *Biocatalyst Design for Stability and Specificity*; Himmel, M. E., Georgiou, G., Eds.; American Chemical Society: Washington, D. C., 1993; Series 516, pp 102-107. Arnold, F.; Haymore, B. L. *Science* **1991,** *252,* 1796-1797.

(18) Kunkel, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1985,** *82,* 488-492.

(19) Casimiro, D. R., Ph.D. Thesis, California Institute of Technology, 1993.

(20) Mayo, Stephen L., Ph.D. Thesis, California Institute of Technology, 1988.

(21) Cutler, R. L.; Pielak, G. J.; Mauk, A. G.; Smith, M. *Prot. Eng.* **1987,** *1,* 95.

(22) Sanger, F.; Nicklen, S.; Coulson, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1977,** *74,* 5463-5467.

(23) Becker, D. M.; Guarente, L. *Methods Enzymol.* **1991,** *194,* 182-187.

(24) Louie, G. V.; Brayer, G. D. *J. Mol. Biol.* **1990,** *214,* 527-555.

(25) Holzschu, D.; Principio, L.; Conklin, K. T.; Hickey, D. R.; Short, J.; Rao, R.; McLendon, G.; Sherman, F. *J. Biol. Chem.* **1987,** *262,* 7125-7131.

(26) Moore, G. R.; Pettigrew, G. W. *Cytochromes c; Evolutionary, Structural and Physicochemical Aspects*; Springer-Verlag: Berlin, 1990; pp 196-198.

(27) Ikeda-Saito, M.; Hori, H.; Andersson, L. A.; Prince, R. C.; Pickering, I. J.; George, G. N.; Sanders, C. R., II; Lutz, R. S.; McKelvey, E. J.; Mattera, R. *J. Biol. Chem.* **1992,** *267,* 22843-22852. Matsuoka, A.; Kobayashi, N.; Shikama, K. *Eur. J. Biochem.* **1992,** *210,* 337-341. Bogumil, R.; Maurus, R.; Hildebrand, D. P.; Brayer, G. D.; Mauk, A. G. *Biochemistry* **1995,** *34,* 10483-10490.

(28) Cheesman, M. R.; Greenwood, C.; Thomson, A. J. *Adv. Inorg. Chem.* **1991,** *36,* 201-255.

Chapter 4

pH-Dependent Equilibria of Fe(III)cytochrome *c* Variants[1]

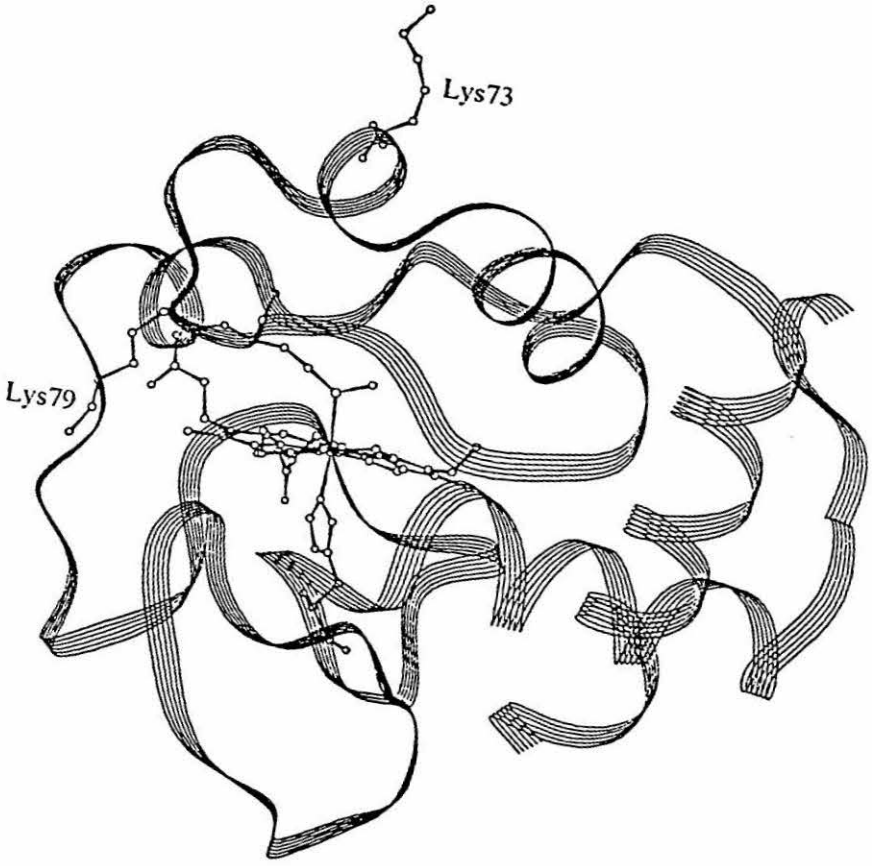
Background

pH-Dependent Equilibria of Fe(III)cytochrome *c*

Fe(III)cytochrome *c* is known to exhibit five major pH-dependent conformational states.[2-4] Spectroscopic studies have shown that the acid denaturation of cyt *c* can be described by a three-state model involving native, molten globule, and acid-unfolded states, with the molten-globule form existing under conditions of high ionic strength.[5] At low pH (< 2.5), the heme crevice opens and the heme undergoes a transition from low-spin to high-spin. It is clear that Met80 is detached from iron coordination at acidic pH values; His18 also may be detached. The conformational equilibria of Fe(III)cyt *c* at alkaline pH values have stimulated particular interest. The "alkaline form" of cyt *c* is formed upon raising the pH to alkaline values ($pK = 8-10$, depending on species).[4] This isomer possesses a six-coordinate, low-spin heme iron that retains native His18 ligation while the Met80 thioether is replaced by a strong-field ligand.[3,4] The identity of this ligand remained controversial for 20 years; while EPR and MCD evidence have suggested that the ε-amino group of a lysine binds the heme iron in the alkaline form,[6] studies of chemically-modified cytochromes have provided conflicting data on this proposal.[7]

^{1}H NMR studies cleared up much of the confusion that arose from the apparent conflict between spectroscopic and chemical-modification results by demonstrating that two alkaline isomers exist.[8] The very similar shifts and linewidths of the hyperfine-shifted heme resonances of the two species suggested that they have similar heme ligands.[8] The observation by ^{1}H NMR of only one alkaline isomer in the Lys79Ala mutant of *S. c.* cyt *c* confirmed that one of the alkaline forms in the wild-type protein has a Lys79-ligated heme.[9] The second alkaline form of *S. c.* cyt *c* has recently been shown through an analogous study to have Lys73 ligated to the heme.[10] This result is surprising, as coordination by Lys73 requires a significant conformational rearrangement of the protein (Figure 4.1).[11]

Figure 4.1 Ribbon diagram of the X-ray crystal structure of *S. c.* WTcyt *c* (PDB identifier 2YCC). The side chains of the two Lys residues (Lys79 and Lys73) proposed to ligate to the heme in the two *S. c.* Fe(III)cyt *c* alkaline isoforms are indicated. While Lys79 is close to the heme and may ligate without a significant conformational change, ligation of Lys73 (or any other Lys) would require an extensive conformational rearrangement of the protein.


Lys73
Lys79

pH Dependence of MetMb Ligation

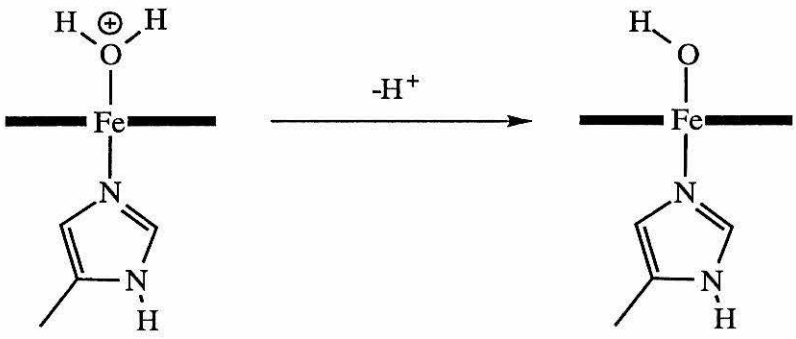
While most studies of ligand-binding to Mb have focused on ferrous derivatives, investigations of the ligand-binding properties of metMb have also shed light on how the heme pocket modulates heme reactivity.[12] MetMb in most cases is high-spin and hexacoordinate with a water molecule at the distal coordination position. For some naturally occurring globins that have hydrophobic residues at the distal His (E7) position (*e.g.*, *Aplysia* Mb, Val(E7)), the oxidized form is pentacoordinate, lacking the heme-bound water.[13] It is believed that the pentacoordinate nature of these Mbs is a result of the hydrophobicity of the E7 residue. Replacement of the distal His in human Mb (His64Leu, His64Val),[14] horse heart Mb (His64Ile),[15] or sperm whale Mb (His64Phe, His64Met)[16] by nonpolar residues leads to the formation of pentacoordinate forms of the oxidized protein, while replacement of His64 with polar residues (His64Gln,[14] His64Asp[16]) retains the coordinated water molecule.

As is observed for Fe(III)cyt *c*, the heme ligation of metMb is pH-dependent. The alkaline transition for Mb (pK 8–9, depending on species) involves deprotonation of the heme-bound water to give a low-spin His-OH^--ligated heme (Figure 4.2).[12] Those Mbs that lack a water molecule at the distal coordination position also undergo an alkaline transition. In the case of *Aplysia* Mb, the data are consistent with the formation of a low-spin His-OH^--ligated heme at alkaline pH values,[17] while in Mb mutants in which the distal His has been replaced by a hydrophobic residue, it has been proposed that the alkaline transition is due to deprotonation of the bound (proximal) His, maintaining a high-spin pentacoordinate heme.[15]

Ala80cyt *c*, because it has a heme capable of binding exogenous ligands, yet maintains the cytochrome fold, is a valuable model system for probing the pH-dependent equilibria of cyt *c* and of Mb. In particular, it allows the study of how heme ligation affects the cytochrome fold as pH is altered and also how the hydrophobic cytochrome interior affects the ligand-binding properties of the heme.

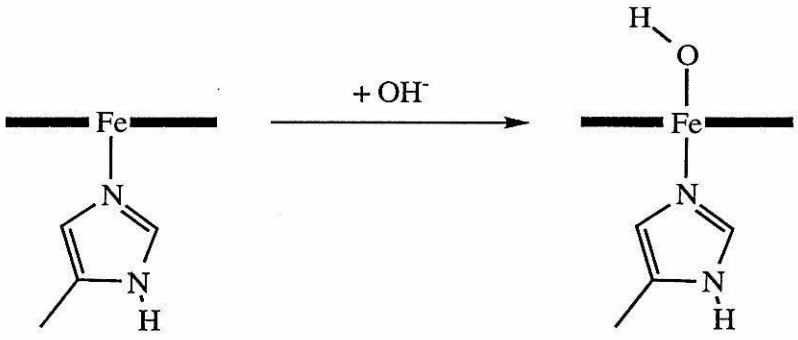
Figure 4.2 Illustration of the alkaline transition as exhibited by different metMbs. (A) Most metMb species undergo an alkaline transition where the heme-bound water is deprotonated to give a low-spin His-OH^--ligated heme. (B) Naturally pentacoordinate Fe(III)Mbs such as *Aplysia* Mb bind OH^- as pH is raised. (C) Pentacoordinate Fe(III)Mb mutants such as His64Val remain high-spin and pentacoordinate through a wide range of pH values, but the proximal His is deprotonated at high pH.

A



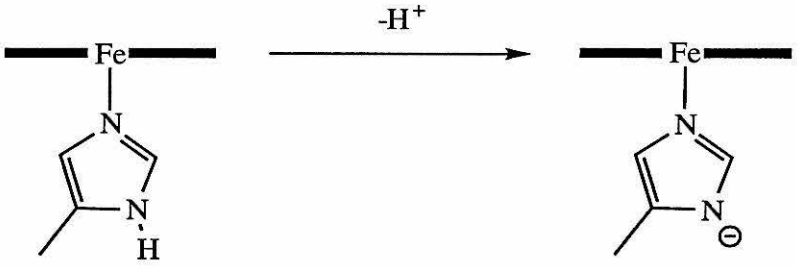
high-spin low-spin

B



high-spin low-spin

C



high-spin high-spin

Paramagnetic Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance spectroscopy has become an invaluable technique for studying the electronic and molecular structure of paramagnetic metalloproteins. The development of instrumentation and new pulse sequences, coupled with a greater understanding of the effects which paramagnetic ions exert on nearby nuclei, has contributed to the successful use of NMR in biological systems containing metal ions with unpaired electrons.[18,19] *S. c.* Ala80cyt *c* is a promising subject for an NMR study due to its relatively small size and its inclusion of the well-characterized heme group. NMR studies of paramagnetic forms of Ala80cyt *c* are expected to lend insight into the questions of how the cytochrome fold regulates the ligand-binding properties of the heme.

Paramagnetic centers exert dramatic effects on the shifts of NMR signals due to nuclei with which they interact. The observed shift of a nucleus in a paramagnetic molecule is the sum of its hypothetical shift in an isostructural diamagnetic complex (δ_{dia}) and its shift due to interactions with the paramagnetic center, the isotropic shift:

$$\delta_{obs} = \delta_{dia} + \delta_{iso} \tag{4.1}$$

The isotropic shift, in turn, has dipolar and contact contributions:

$$\delta_{iso} = \delta_{dip} + \delta_{con} \tag{4.2}$$

The dipolar shift results from a through-space interaction between the magnetic moments of the electron and nucleus as modulated by molecular rotation. This effect can be approximated by a function[20]

$$\left(\frac{\Delta\nu}{\nu_0}\right) = \left[\frac{1}{4\pi}\frac{1}{3N_A}\chi_{zz} - \frac{1}{2}(\chi_{xx}+\chi_{yy})\frac{3\cos^2\theta - 1}{\mathbf{r}^3}\right] - \frac{1}{4\pi}\frac{1}{3N_A}(\chi_{xx}-\chi_{yy})\frac{\sin^2\theta\cos 2\phi}{\mathbf{r}^3} \tag{4.3}$$

where $\Delta\nu$ is the averaged hyperfine shift, ν_0 is the resonance frequency of a reference nucleus at the magnetic field of the experiment, N_A is Avagadro's number, χ_{xx}, χ_{yy}, and χ_{zz} are the three principal components of the magnetic susceptibility tensor, and $\mathbf{r}$, θ, and ϕ are polar coordinates relating the position of the resonating nucleus within the axis

frame determined by the main directions of the χ tensor. In solution, molecular tumbling results in an isotropic contribution of this factor which is then called the pseudocontact shift. The contact shift results from the presence of unpaired electron spin density at the nucleus in question, usually through spin polarization effects.[18] In other words, nuclei with a through-bond interaction with a paramagnet will experience a contact contribution to the hyperfine shift as described by a function[21]

$$\left(\frac{\Delta\nu}{\nu_0}\right) = \frac{2\pi A_c}{h}\, S(S+1)\, \frac{\mathbf{g}\mu_B B_0}{3\gamma_N kT} \tag{4.4}$$

which depends on the Fermi contact coupling constant A_c, Planck's constant h, the spin of the metal ion S, the **g** tensor, the Bohr magneton μ_B, the applied magnetic field B_0, the magnetogyric ratio for the nucleus γ_N, the Boltzmann constant k, and the absolute temperature T.

As unpaired electrons provide an efficient relaxation pathway for nuclei to which they are coupled, the result is significant broadening of NMR lines in paramagnetic systems. As in the case of the hyperfine shift, there is both a dipolar[22] and a contact[23] contribution to relaxation, which in the fast-motion limit is given by equations 4.5 and 4.6, respectively.

$$T_{1M}^{-1} = T_{2M}^{-1} = \frac{4}{3}\left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_N^2 g_e^2 \mu_B^2 S(S+1)}{r^6}\, T_{1e} \tag{4.5}$$

$$T_{1M}^{-1} = T_{2M}^{-1} = \frac{2}{3}\, S(S+1)\left(\frac{2\pi A_c}{h}\right)^2 \tau_s \tag{4.6}$$

where T_{1e} is the longitudnal electronic relaxation time and τ_s is the electronic relaxation time. In addition, a third mechanism, Curie spin relaxation, introduces the contribution of the external magnetic field to linewidths. Often, linewidths of hyperfine-shifted signals of slowly tumbling molecules decrease with lower magnetic field and in fact lines undetectable at 600 MHz may be well-resolved at 200 MHz.[18]

Because of the extreme effects exerted by unpaired electrons on nearby nuclei, the electronic properties of the paramagnetic ion in the system are of great importance in determining the suitability of a paramagnetic system for NMR studies. While organic free radicals and some metal ions such as Gd(III), Cu(II), and Mn(II) have long electronic relaxation times, leading to extreme broadening of NMR signals, other ions (including low-spin Fe(III)) have favorable electronic properties for paramagnetic NMR. The short τ_s values displayed by these ions result in relatively long T_1's for nearby nuclei and, therefore, sharp NMR lines (Table 4.1).[18]

Table 4.1

Electronic relaxation times of selected metal ions[18]

Metal Ion	τ_s (seconds)
Mn(III)	10^{-10}–10^{-11}
Mn(II)	10^{-8}–10^{-9}
Fe(III) (H.S.)[a]	10^{-10}–10^{-11}
Fe(III) (L.S.)	10^{-11}–10^{-12}
Fe(II) (H.S.)	10^{-12}
Co(II) (H.S.)	10^{-11}–10^{-12}
Co(II) (L.S.)	10^{-9}–10^{-10}
Cu(II)	1–3 x 10^{-9}
Gd(III)	10^{-8}–10^{-9}

[a]H.S. denotes high-spin; L.S. denotes low-spin.

Although the effect of a paramagnetic center on nuclear spin is complex in theory, for biomolecules, the resulting NMR spectra may actually be simplified compared to those of diamagnetic biomolecules. This is due to the large range of chemical shifts found for protons close in space or bound to the paramagnetic center. The result of this effect in paramagnetic metalloproteins is that the resonances of active-site nuclei are well resolved from what is termed the diamagnetic envelope of overlapping peaks observed

between 0 and 8 ppm (in a ^{1}H NMR spectrum of a protein). Although some paramagnetic centers exert an extreme broadening effect on the NMR signals of nearby nuclei, the paramagnetic forms of heme proteins often have favorable electronic properties which result in shifting resonances without deleterious broadening.

Materials and Methods

General

Semisynthetic horse heart Ala80cyt *c* and Phe67Ala80cyt *c* were prepared and purified as described in Chapter 2.[18] *Saccharomyces cerevisiae* Gln39Ala80Ser102-iso-1-cyt *c* (*S. c.* Ala80cyt *c*) was expressed and purified as described in Chapter 3.[19] Protein handling and reagents used were discussed in Chapter 2. *S. c.* WTcyt *c* is *S. c.* Ser102cyt *c*. Sperm whale Mb was obtained from Sigma.

Absorption Spectroscopy

UV/vis absorption spectroscopy was performed on protein samples in 25 mM NaP_i buffer in a 1.000-cm quartz cuvette. The HP8452 diode-array spectrometer was used for all absorption measurements in pH titrations because it allows the collection of data quickly; this is particularly important at extreme pH values when sample pH is not stable over long periods of time. The pH of UV/vis samples was adjusted by adding sub-micromolar to micromolar volumes of 0.5–2 M NaOH and of 3–30% H_3PO_4 immediately followed by gentle pipetting to mix. Sample pH was measured using a Beckman model Φ32 pH meter and an Ingold microelectrode standardized with VWR buffer standards.

^{1}H NMR Spectroscopy

The ^{1}H NMR samples were prepared by dissolving lyophilized protein in 50 mM phosphate buffer at pH 7.0 to give solutions 1 mM in protein. The pH of the NMR

samples was adjusted by addition of small volumes of concentrated solutions of NaOH or H_3PO_4. The pH was measured (uncorrected for the isotope effect) with an Orion model 720 pH meter and a Microelectrodes Inc. model MI-410 microcombination pH probe.

The ^1H NMR spectra were recorded on a Bruker MSL 200 spectrometer using a superWEFT (water-eliminated Fourier transform) pulse sequence[24] with a recycle delay of 200 to 250 ms.

Nonselective T_1 measurements were taken at 200 MHz with the inversion recovery pulse sequence[25] with delay times between subsequent pulses from 2048 to 0.05 ms and a recycle delay of 2048 ms. The peak intensity was fitted to a single-exponential function to obtain the T_1 value. The recovery of magnetization after a nonselective excitation is expected to be essentially exponential for nuclei near the iron, for which relaxation by the unpaired electron dominates over cross relaxation.[26-28] When the paramagnetic contribution is sizable but not dominant as in the experiments reported here, only the initial part of the magnetization recovery of a nonselective experiment is expected to fit a single exponential. However, within the limitations of these experiments and for the protein derivatives studied, the same T_1 values were found by fitting either the initial part of the magnetization recovery or the whole magnetization recovery curve to a single exponential (*vide infra*).

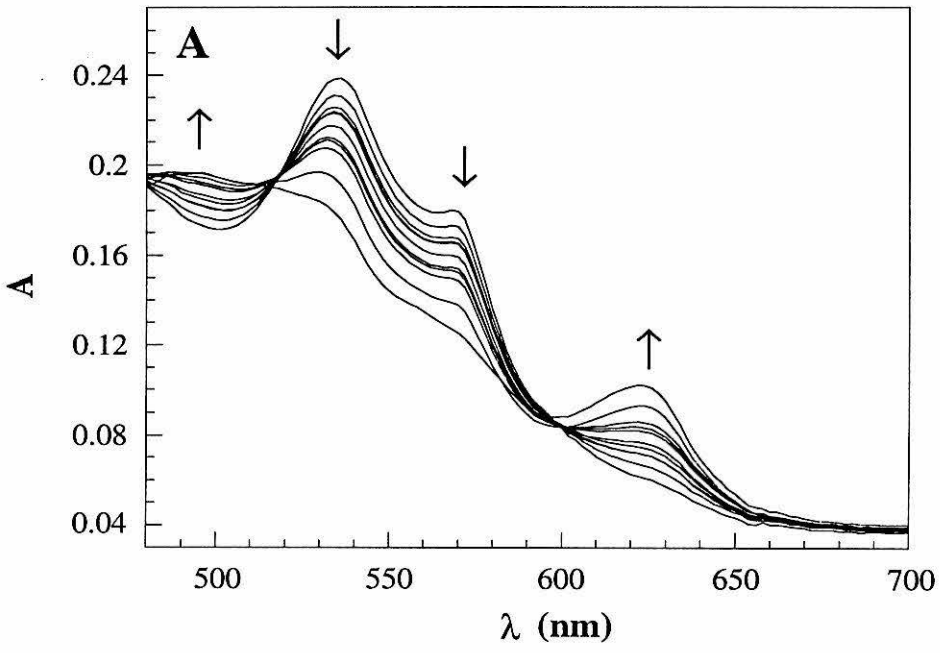
Results and Discussion

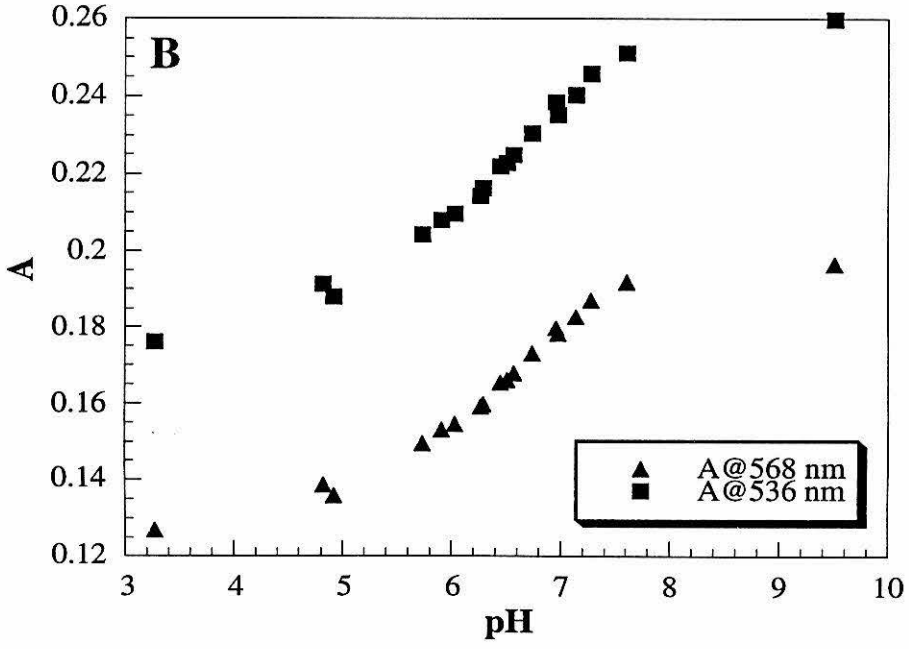
Absorption Spectroscopy

S. c. Fe(III)Ala80cyt c

The visible absorption spectrum of *S. c.* Fe(III)Ala80cyt *c* undergoes a two-state transition as pH is raised from pH 3 to 10. The α, β-region of the absorption spectrum in this pH range is shown in Figure 4.3A. Plotting the absorbance at a given wavelength vs. pH results in a titration curve with a single inflection point at pH 6.6, shown in Figure

Figure 4.3 (A) pH dependence of the visible absorption spectrum of *S. c.* Fe(III)Ala80cyt *c*. The sample was 25 μM in 25 mM NaP_i. The arrows indicate the direction of the spectral changes as pH is lowered. The spectral changes are characteristic of a transition from low-spin to high-spin. (B) Plot of absorbance vs. pH at selected wavelengths. The midpoint transition is at pH 6.6; this is presumed to be the pK_a for the deprotonation of the heme-bound water.


A
0.24
0.2
0.16
0.12
0.08
0.04
A
500
550
600
650
700
λ (nm)


B
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
A
3
4
5
6
7
8
9
10
pH
A@568 nm
A@536 nm

4.3B. The similarity of the UV/vis spectra of the mutant at low and high pH with those of Mb under conditions where it is known to have water (pH 7) or hydroxide (pH 10) bound to the heme[12] suggests that the mutant has His-H_2O-ligation below pH 6 and His-OH^--ligation above pH 7 (Chapter 3). The spectral changes are reversible in this pH range.

The pK_a of 6.6 is 2–3 units lower than the pK_a measured for the same transition in hexacoordinate Mbs, but is similar to that observed for peroxidase mutants that bind water.[29] One possible explanation for the low pK_a of the heme-bound water in Fe(III)Ala80 is that the extremely hydrophobic cyt *c* pocket stabilizes the neutral hydroxide-bound species over the charged water-ligated species. An additional possibility is that hydrogen-bonding from the distal pocket residue Tyr67 preferentially stabilizes the hydroxy species (Figure 4.4).

Horse Heart Fe(III)Ala80cyt c

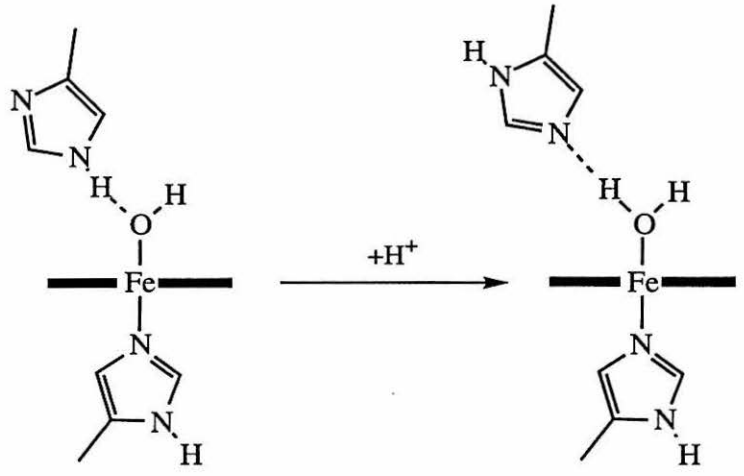
The UV/vis absorption spectrum of semisynthetic Fe(III)Ala80 displays pH-dependent changes (Figure 4.5A) highly analogous to those exhibited by *S. c.* Fe(III)Ala80. Characterization of semisynthetic Fe(III)Ala80 by EPR and MCD spectroscopy (Chapter 2) confirms that ionization of the heme-bound water/hydroxide is responsible for the spectral changes, which are reversible between pH 4 and 10. The midpoint pH of this transition for semisynthetic Ala80 is 6.3.

Horse Heart Fe(III)Phe67Ala80cyt c

Semisynthetic Phe67Ala80 undergoes an alkaline transition with a pK of 6.3 (Figure 4.6). The pH dependence of the UV/vis spectrum is very similar to that observed for Fe(III)Ala80 from both sources, again suggesting that a heme-bound water is ionizing in this transition. However, this mutant has not been characterized by MCD or EPR spectroscopy, making this ligation assignment tentative at this time.

Figure 4.4 Illustration of interactions of (A) a distal His side chain and (B) a distal Tyr side chain with a water/hydroxide species bound to a heme. While the His is able to tautomerize to accommodate the water/hydroxide transformation, the Tyr may need to deprotonate to accommodate protonation of a heme-bound water. The presence of a group with a high (~ 10 for Tyr in a peptide) pK_a may account for the low pK_a exhibited by the heme-bound water in Fe(III)Ala80cyts *c*.

A



B

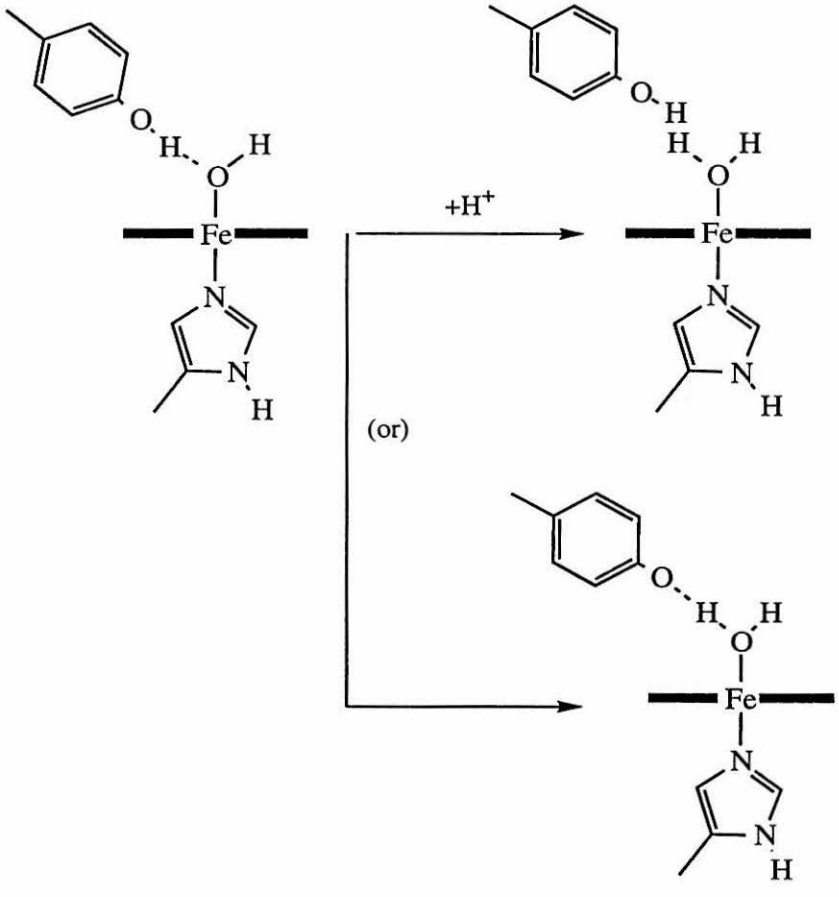
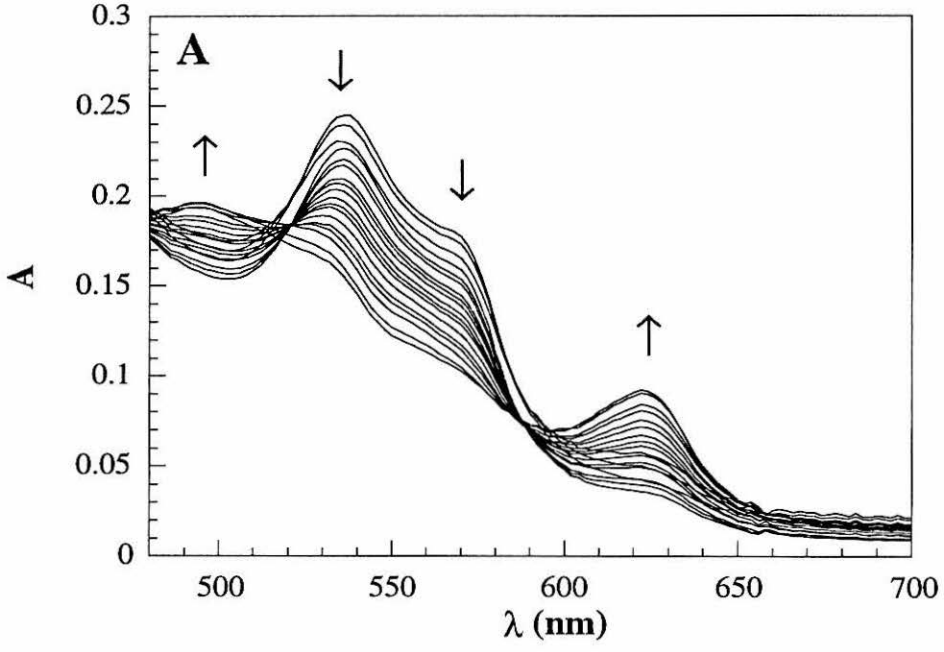


Figure 4.5 (A) pH dependence of the visible absorption spectrum of h. h. Fe(III)Ala80cyt *c*. The sample was 26 μM in 25 mM NaP_i. The arrows indicate the direction of spectral changes as pH is lowered. (B) Plot of absorbance vs. pH at selected wavelengths. The data indicate that the heme-bound water has a pK_a of 6.3.


A
A
λ (nm)
0.3
0.25
0.2
0.15
0.1
0.05
0
500
550
600
650
700

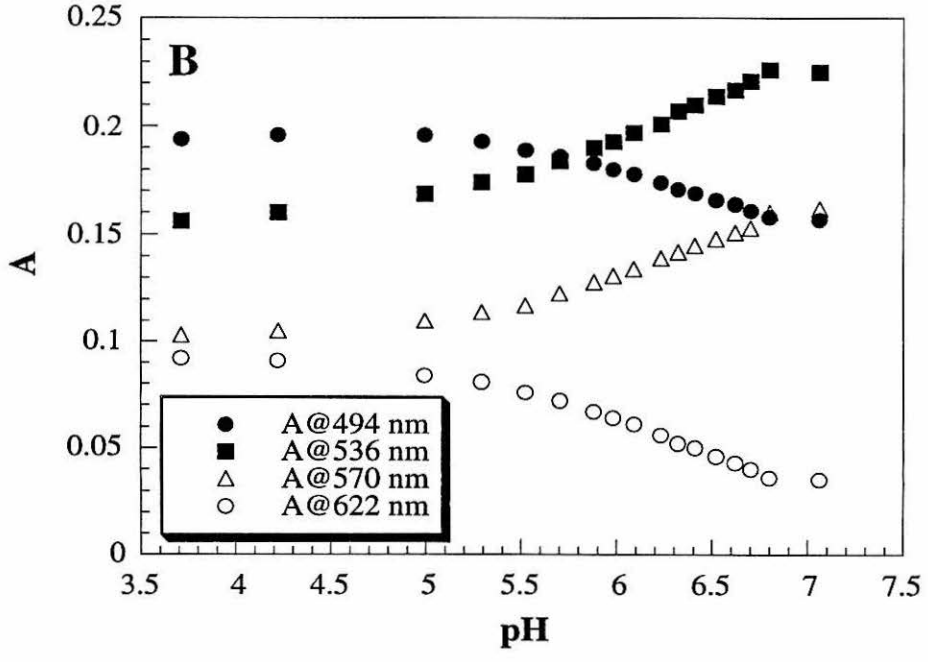
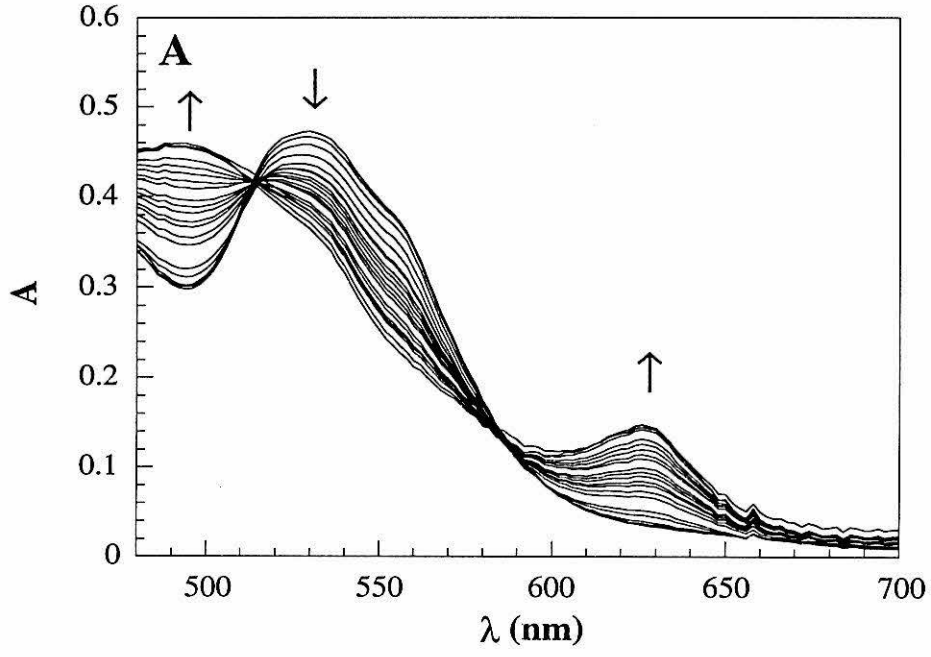
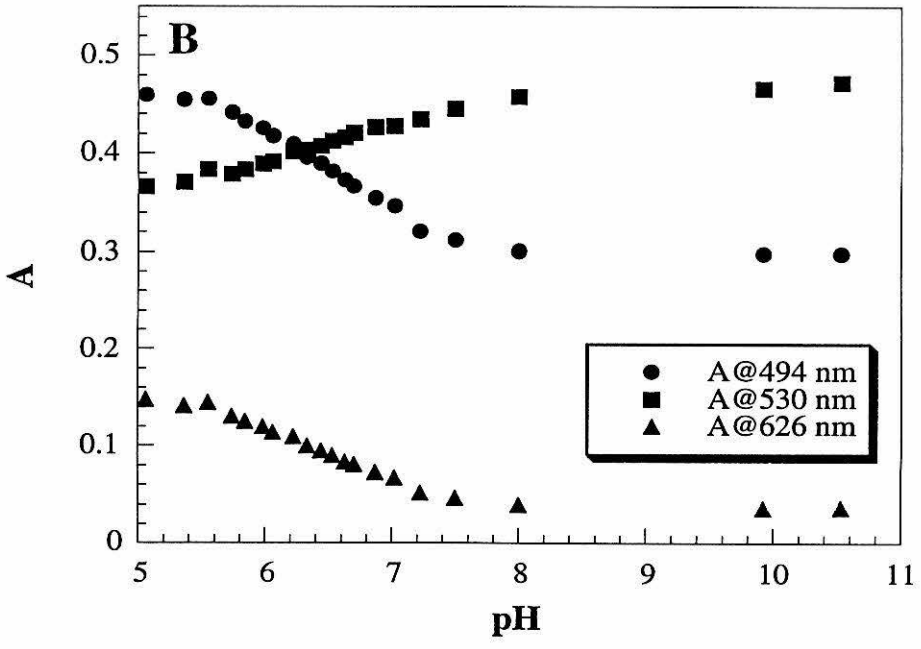

B
A
pH
A@494 nm
A@536 nm
A@570 nm
A@622 nm
0.25
0.2
0.15
0.1
0.05
0
3.5
4
4.5
5
5.5
6
6.5
7
7.5

Figure 4.6 (A) pH dependence of the visible absorption spectrum of h. h. Fe(III)Phe67Ala80cyt *c*. The sample was 50 mM in 25 mM NaP_i. The arrows indicate the direction of spectral changes as pH is lowered. (B) Plot of absorbance vs. pH at selected wavelengths. The transition midpoint is at pH 6.6. The changes, which resemble those seen for the Ala80cyt *c* mutants, may be due to protonation of a heme-bound hydroxide.


A
A
λ (nm)
0.6
0.5
0.4
0.3
0.2
0.1
0
500
550
600
650
700


B
A
pH
A@494 nm
A@530 nm
A@626 nm
0.5
0.4
0.3
0.2
0.1
0
5
6
7
8
9
10
11

Paramagnetic ^{1}H NMR Spectroscopy of *S. c.* Fe(III)Ala80cyt *c* and *S. c.* Fe(III)WTcyt *c*

Spectra at Intermediate pH (4–8)

The ^{1}H NMR spectrum of *S. c.* Fe(III)Ala80 undergoes dramatic changes (Figure 4.7A–C) as the sample pH is increased from 4 to 7. The most significant change is a transition from high-spin (species HS3, see Table 4.2) to low-spin (LS3), which is slow on the 200 MHz ^{1}H NMR timescale. Evaluation of the NMR signal intensities in this pH range gives a p*K* of ~ 6 for this transition, which correlates well with the value of 6.6 obtained from the pH dependence of the electronic absorption spectra. The HS3 species is characterized by broad resonances over a wide spectral range (in particular, around 60 ppm), as expected for heme methyls coupled to a high-spin iron.[18,19,30] The spectrum of the HS3 species, which is dominant between pH 4 and 5, undergoes minor changes with pH (Figure 4.7B,C). Species LS3 is dominant between pH 7 and 8; it is characterized by hyperfine-shifted signals in the 30–5 ppm region (Figure 4.7A).

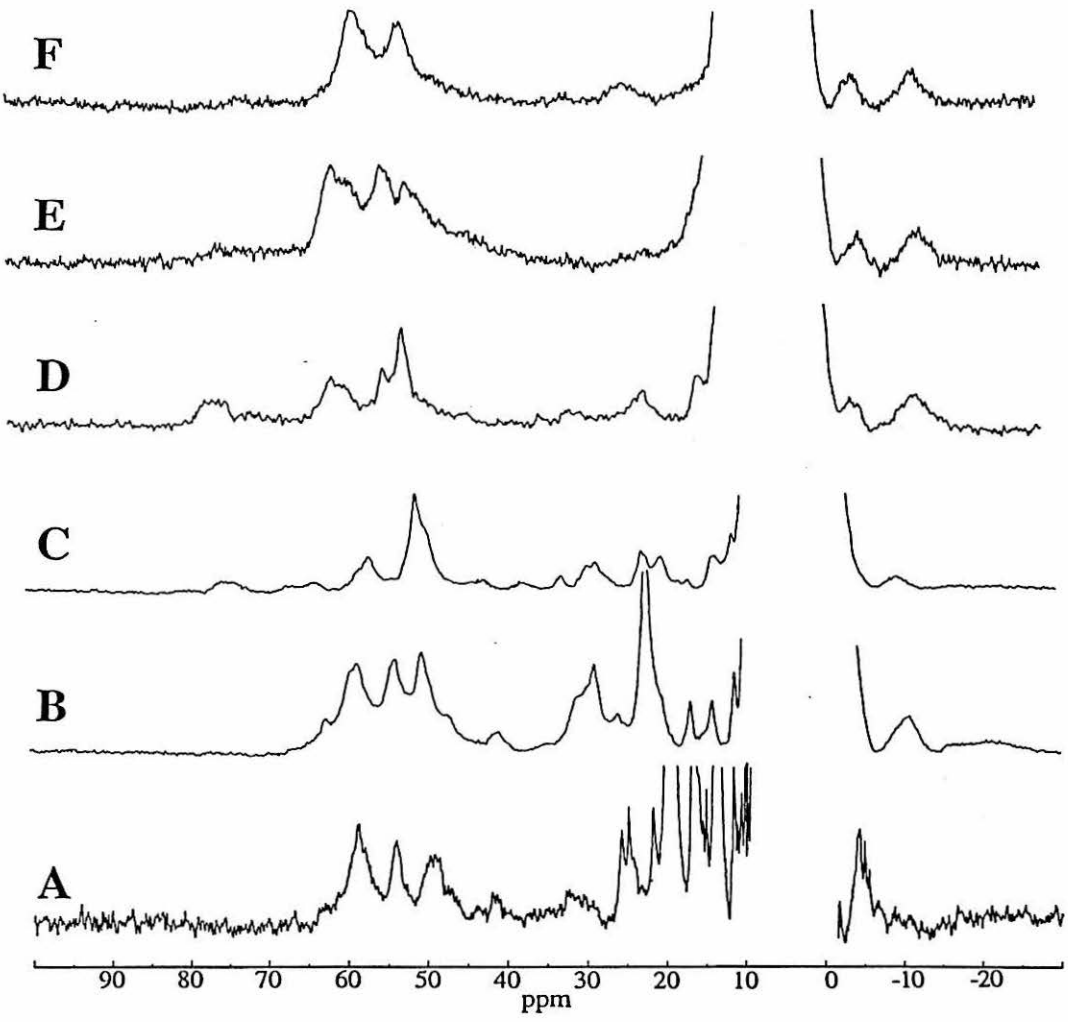
Table 4.2

S. c. WTcyt c and S. c. Ala80cyt c species at different pH values[a]

	S. c. WTcyt *c*		*S. c.* Ala80cyt *c*	
pH	heme methyl shift range (ppm)	species	Heme methyl shift range (ppm)	species
2.2	60–45	HS1	60–45	HS1
3.0	80–40	HS2	80–45	HS2
4.0	35–30	LS1	80–45	HS2
6.0	35–30	LS1	70–40 + 35–13	HS3 + LS3
7.0	35–30	LS1	20–13	LS3
9.0	35–30 + 25–10	LS1 + LS2(a+b)	20–13	LS3
11.0	25–10	LS2(a+b)	25–10	LS2(a)

[a]HS indicates a high-spin species, LS a low-spin one.

Figure 4.7 200 MHz ^{1}H NMR spectra of *S. c.* Fe(III)Ala80cyt *c*. The sample was 1 mM in 50 mM NaP_i in D_2O at (A) pH = 7.0, T = 296 K; (B) pH = 5.0, T = 296 K; (C) pH = 4.0, T = 296 K; (D) pH = 3.7, T = 296 K; (E) pH = 2.2, T = 296 K; (F) pH = 2.2, T = 306 K. The vertical scale in A is 8x that of B; the B–F scales are the same.


F
E
D
C
B
A
90
80
70
60
50
40
30
20
10
0
-10
-20
ppm

The nonselective T_1 values of the heme methyl protons of *S. c.* Ala80cyt *c* derivatives compared to those of other heme proteins are reported in Table 4.3. When electron-nuclear coupling dominates nuclear relaxation, the experimental magnetization recovery is expected to be exponential (Figure 4.8).[26-28] This is the case for the T_1 values reported here. The experimental T_1 values of species LS3 are shorter than 10 ms (Table 4.3), that is, they are dominated by the coupling with the unpaired electron. Under these circumstances, the T_1 values depend on the sixth power of the proton-iron distance and on the electron correlation time τ_s. The τ_s value, in turn, is related to the iron ligand field. A comparison of T_1 values of heme methyls among related heme proteins thus provides meaningful information in relative electronic relaxation times because the distances between the methyls and the iron in the heme are constant and the only variable is τ_s. The LS3 species of Ala80cyt *c* has T_1 values of the same order of magnitude as the alkaline form of metMb, which is believed to contain a HO-Fe(III) species.[12] It is likely that the weak OH^- ligand causes strong axial distortions and long electronic relaxation times.[26,31] Note that the T_1 values of LS3 are much shorter than those of hemes with two

Table 4.3

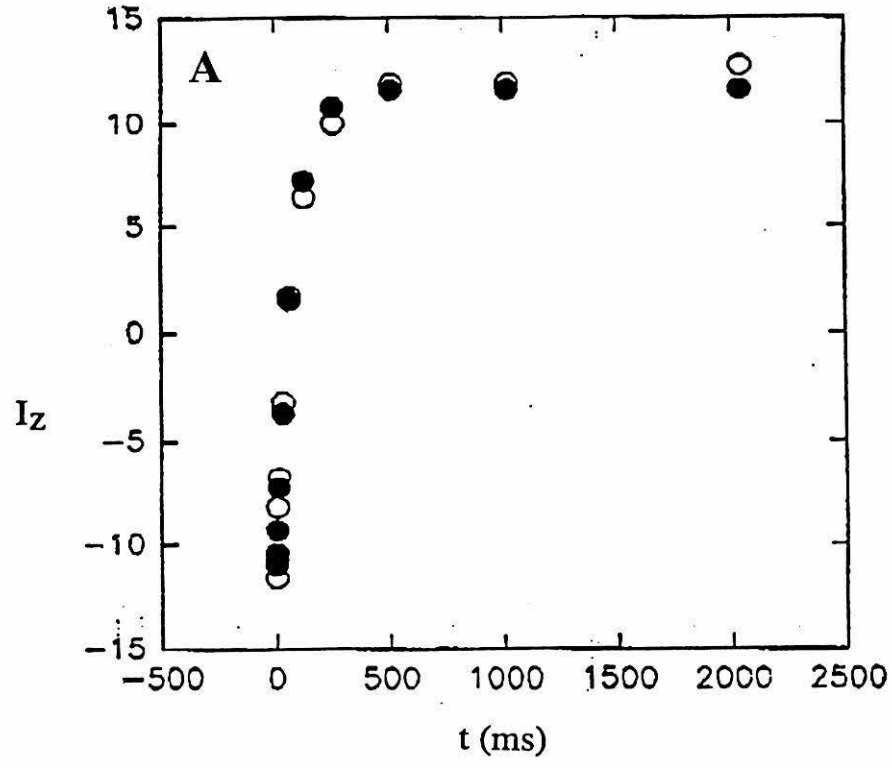
Nonselective T_1 values for heme-methyl ^1H resonances in selected heme proteins[a]

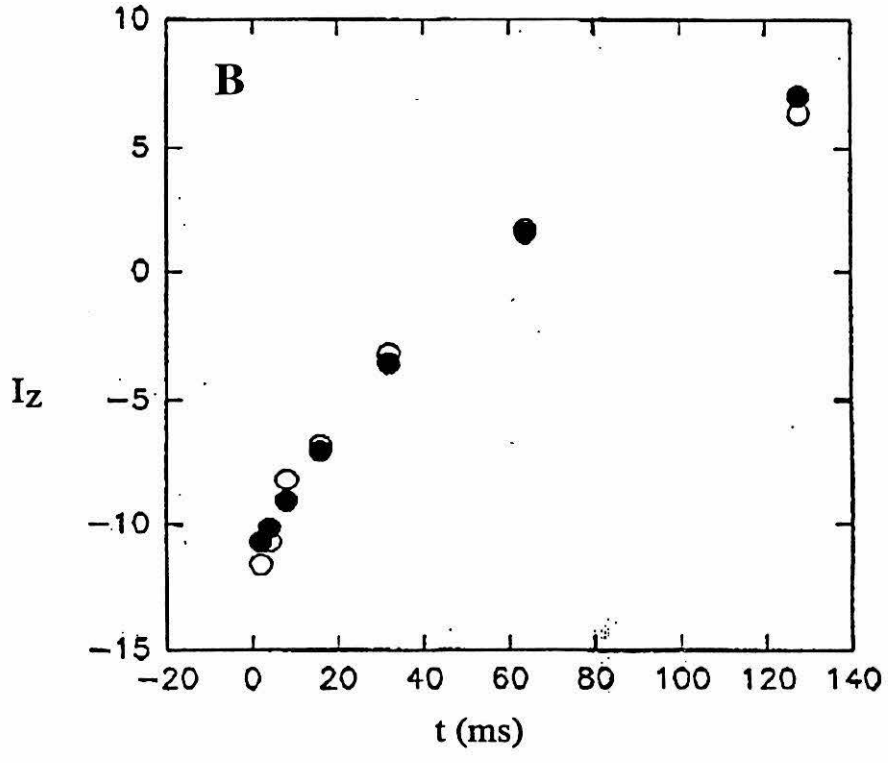
Protein	pH	Species	Heme methyl T_1's
Fe(III)WTcyt *c*	7	LS1	45–50 ms
Fe(III)WTcyt *c*	11	LS2(a+b)	50–85 ms
Fe(III)Ala80cyt *c*	7	LS2a	5.5–9 ms
Fe(III)Ala80cyt *c*	11	LS2a	55–95 ms
HO-Fe(III)Mb	10.4	His/OH^- heme ligation	1–2 ms
CN-Fe(III)Ala80cyt *c*	7	His/CN^- heme ligation	73–115 ms
CN-Fe(III)Pxs[b]		His-/CN^- heme ligation	40–90 ms

[a]Measured at 200 MHz.

[b]Pxs denotes peroxidases. Data on Pxs-CN^- are from reference 32.

Figure 4.8 T_1 plots for the heme 8-CH_3 protons of the cyanide derivative of *S. c.* Fe(III)Ala80cyt *c*. (A) Recovery of the heme 8-CH_3 magnetization after nonselective irradiation (open circles) fit to a single exponential (closed circles). The inversion recovery pulse sequence was used. The data indicate a T_1 value of 76 ± 5.0 (3σ) ms. (B) Recovery of the heme 8-CH_3 magnetization after nonselective irradiation (open circles) fit to a single exponential (closed circles), discarding the last four experimental values. The data indicate a T_1 value of 76 ± 4.0 (3σ) ms.


A
15
10
5
0
−5
−10
−15
−500
0
500
1000
1500
2000
2500
Iz
t (ms)


B
10
5
0
−5
−10
−15
−20
0
20
40
60
80
100
120
140
Iz
t (ms)

strong-field ligands, such as CN-His and His-Met-ligated species (Table 4.3),[32] consistent with hydroxide ligation. These results support the assignment of axial His-OH^- ligation for Ala80cyt *c* at pH 7.

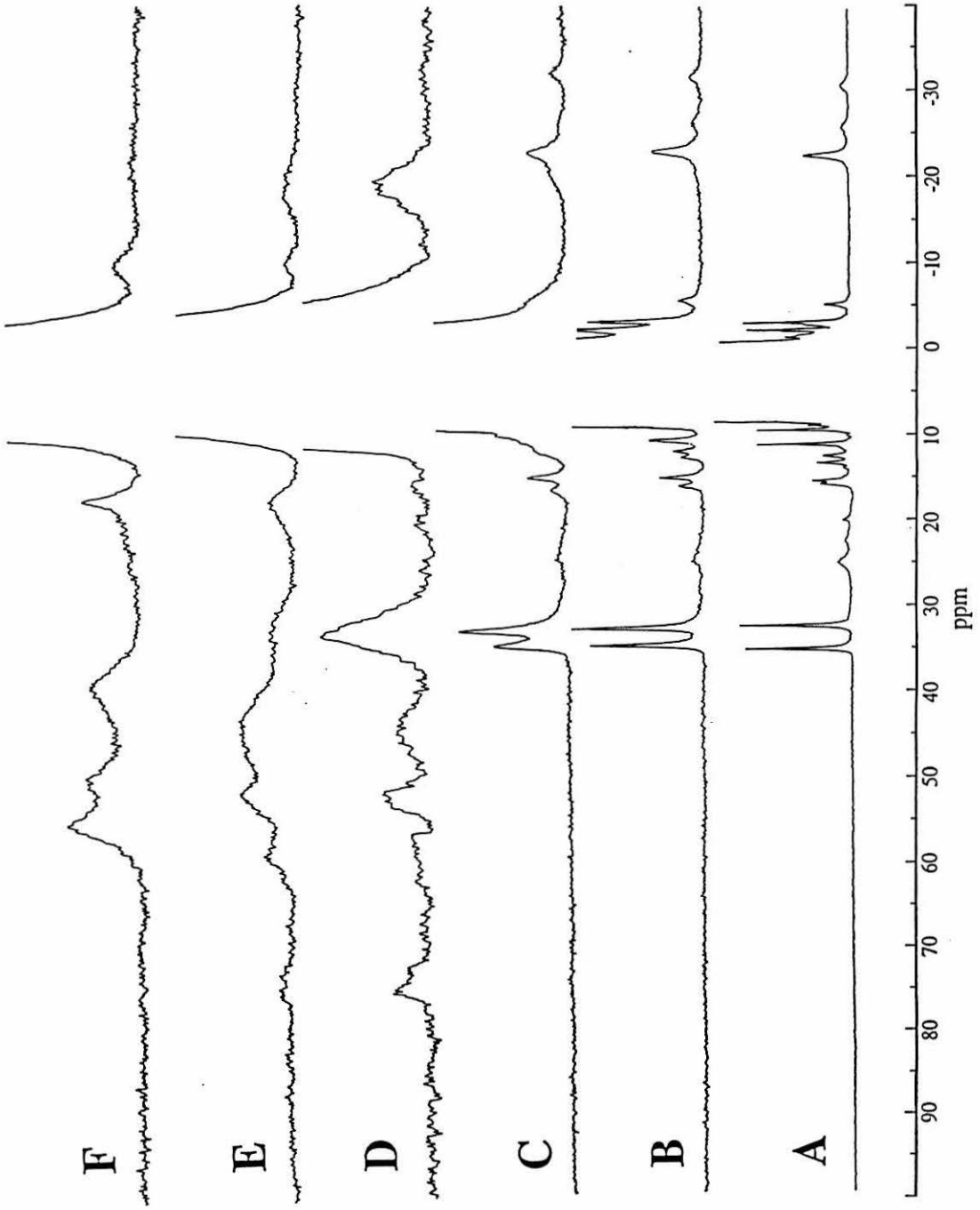
The ^1H NMR spectra of Fe(III)WTcyt *c* between pH 4–8 indicate the presence of a single low-spin Fe(III) species (LS1). The spectra (Figure 4.9A–C) show two downfield-shifted resonances of intensity 3 that have previously been assigned to heme methyls.[33] The single-proton signals in the 25–8 ppm region are due to protons of other heme substituents and to the axial ligands.[33] Resonances of the axial Met and His protons are also present upfield in the -23 to -32 ppm range.[33] The shifts and T_1 values (Table 4.3) are consistent with the values expected for a system containing a hexacoordinate low-spin Fe(III), where both axial ligands are protein side chains.[18,19] The small pH dependence of the shifts of the hyperfine-shifted resonances in the 4–8 pH range is also consistent with that previously reported.[8]

Spectra at Low pH (2–4)

As the pH is decreased from 5 to 4, new broad resonances appear in the ^1H NMR spectrum of Ala80cyt *c* in the 80–40 ppm range (species HS2) while the signals of species HS3 decrease in intensity (Figure 4.7B–D). Species HS2 and HS3 are in slow exchange on the 200 MHz ^1H NMR timescale. At pH 2.2 and 296 K (Figure 4.7E), the NMR spectrum of the mutant indicates another high-spin species (HS1) with four signals in the 60–45 ppm region. This species also is in slow exchange with HS2 and HS3. It appears that there are at least three high-spin species in equilibrium, with one, HS2, appearing and disappearing within three pH units. Its existence therefore depends on more than one ionization. Temperature exerts a strong influence on the distribution of the species (Figure 4.7E, F).

When the pH of *S. c.* Fe(III)WTcyt *c* is lowered below 4, the hyperfine-shifted signals broaden and other resonances, characterized by larger linewidths, appear in the

Figure 4.9 200 MHz ^{1}H NMR spectra of *S. c.* Fe(III)WTcyt *c*. The 1-mM protein solutions were in 50 mM NaP_i in D_2O at (A) pH = 8.2, T = 296 K; (B) pH = 4.5, T = 296 K; (C) pH = 4.0, T = 296 K; (D) pH = 3.0, T = 296 K; (E) pH = 2.2 T = 296 K; (F) pH = 2.2, T = 306 K. Reported pH values are not corrected for the isotope effect.


F
E
D
C
B
A
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
ppm

80-40 ppm region (Figure 4.9C, D). This behavior suggests quasi-slow exchange (on the 200 MHz ^{1}H NMR timescale) between the neutral low-spin species (LS1) and a high-spin one (HS2). From the shift differences between the resonances of these two species, it is possible to estimate that the exchange rate is on the order of 2–4 x 10^4 s^{-1}, which is within the minimum and maximum difference in chemical shifts between methyl signals of the two species. The ^{1}H NMR spectrum of this species is very similar to that of the mutant HS2 species. The formation of HS2 could be due to bond breaking between the iron and one or both axial ligands, producing a high-spin Fe(III) form. A further decrease in pH (pH 2.2, Figure 4.9E) produces another species (HS1) in slow exchange with HS2. HS1 is characterized by broad resonances in the 60–50 ppm region, and is very similar to the analogous species seen for the mutant. Temperature affects both the shifts of the hyperfine-shifted signals (a decrease in temperature producing an increase in shift) and the distribution of the species (Figure 4.9E, F).

Spectroscopic studies have shown that the acid denaturation of WTcyt c[2-5] involves at least two forms (molten globule and unfolded).[5] Here, in essentially the same pH range, two high-spin species in both variants are observed. The high-spin nature of the low-pH forms of WTcyt *c* (HS1 and HS2) indicates that at least one of the two protein axial ligands is not coordinated to the iron. The strong similarities between the spectra of WTcyt *c* and those of Ala80cyt *c* recorded at low pH suggest that these proteins have the same heme coordination under these conditions, and therefore Met80 is not coordinated at low pH in WTcyt *c*. It is possible that the HS2 species of the two proteins, which appear as the pH is lowered below 4, still have the His coordinated to the iron and that its detachment is associated with the HS2 → HS1 transition around pH 2. For the HS2 species, the shifts in the diamagnetic region of the NMR spectra show dispersion consistent with the presence of some secondary structure.

Spectra at High pH (8–11)

The alkaline transition of WTcyt *c* is manifested in the NMR spectrum by new resonances that appear in the 25–10 ppm region.[8,9] Consistent with the pK of 8.5 for this equilibrium determined by electronic absorption spectroscopy,[34] the alkaline transition is almost complete at pH 10 (Figure 4.10A–C). The T_1 values of these resonances are only slightly longer than those of the neutral pH form LS1 (Table 4.3). When the pH is increased to 11, the intensities of some of these signals (indicated by x in Figure 4.10C, D) in the 25–10 ppm range decrease, suggesting the existence of two species with very similar chemical shifts and T_1 values (designated LS2(a+b)). The equilibrium between the two alkaline species is sensitive to temperature, with an increase in temperature mimicking a decrease in pH (Figure 4.10C–E). The shift data at high pH presented here are consistent with literature values.[9] The T_1 values have not previously been reported.

The spectrum of Fe(III)Ala80 is essentially pH insensitive between pH 7 and 9 (Figure 4.11A, B), but changes drastically above pH 9, with a pK of ~ 10, as determined from the pH dependence of the intensity of the hyperfine-shifted signals. The spectrum recorded at pH 11 (Figure 4.11D) closely resembles that of the dominant species of WTcyt *c* at the same pH (Figure 4.10D). In addition, the heme methyls of both proteins exhibit the same T_1 values (Table 4.2), and these values are also very similar to those of the other low-spin form (LS1) of WTcyt *c*. The T_1 values of the heme methyls of Ala80cyt *c* LS2 are significantly longer than those of Ala80cyt *c* at pH 7, suggesting that a stronger-field ligand is coordinated at high pH. In the mutant, temperature affects the value of the pK for the alkaline transition, suggesting that a conformational change is involved.

The two alkaline forms of WTcyt *c* observed by ^{1}H NMR have been attributed to ligation of two different lysine side chains, one of which is Lys79 (Figure 4.1).[8,9] The

Figure 4.10 200 MHz ^{1}H NMR spectra of *S. c.* Fe(III)WTcyt *c*. The spectra were recorded in 50 mM NaP_i in D_2O at (A) pH = 8.2, T = 296 K; (B) pH = 9.1, T = 296 K; (C) pH = 10.0, T =296 K; (D) pH = 11.0, T = 296 K; (E) pH = 10.0, T = 306 K. pH values are not corrected for the isotope effect. The x's indicate the signals that decrease in intensity as pH is raised above 10.

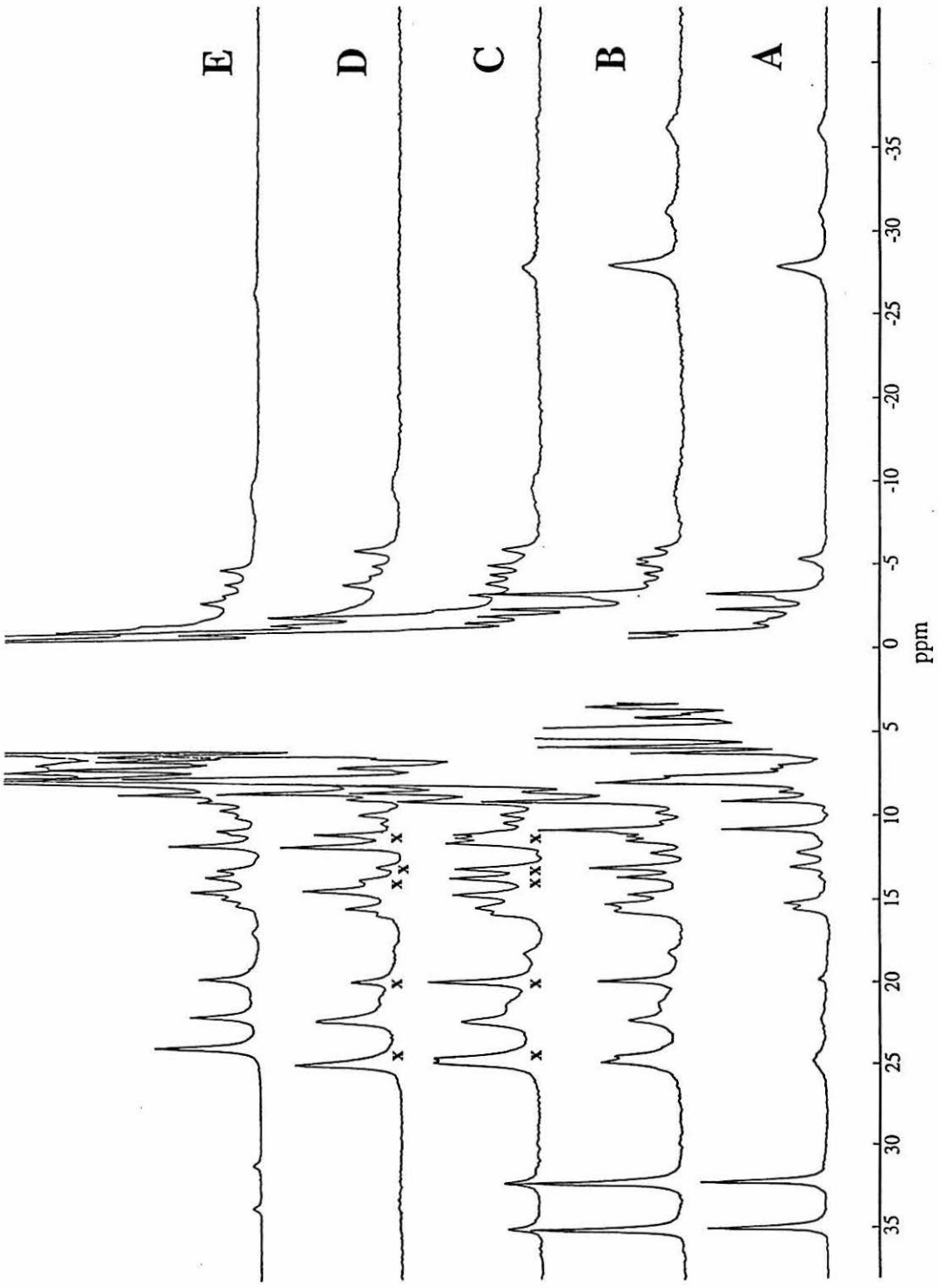
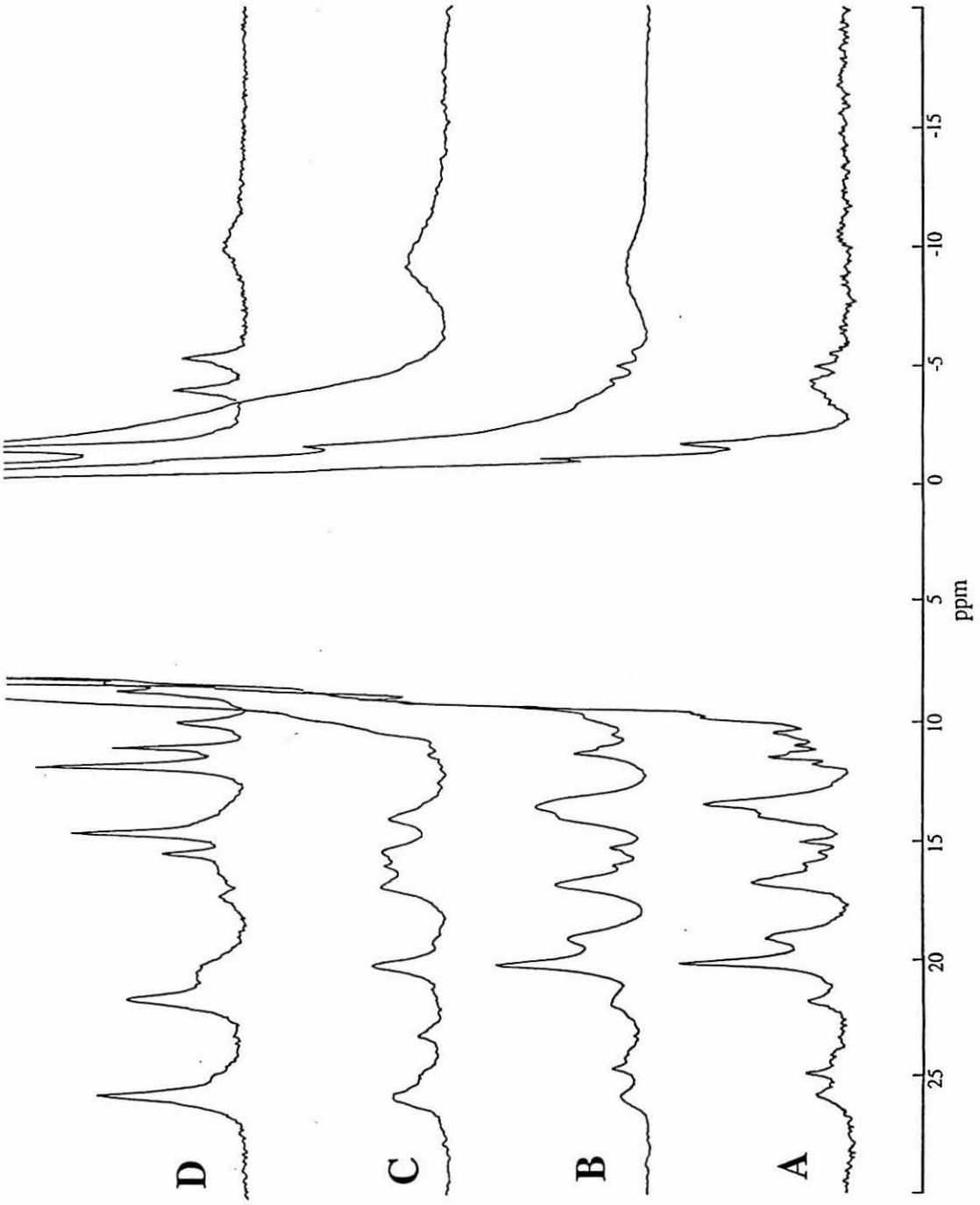

E
D
C
B
A
x
x
x
x
x
x
x
x
x
35
30
25
20
15
10
5
0
-5
-10
-20
-25
-30
-35
ppm

Figure 4.11 200 MHz ^{1}H NMR spectra of 1-mM *S. c.* Fe(III)Ala80cyt *c*. The spectra were recorded in 50 mM NaP_i in D_2O at 296 K and: (A) pH = 7.0, (B) pH = 9.0, (C) pH = 10.0; (D) pH = 11.0 (uncorrected pH values).


D
C
B
A
25
20
15
10
5
0
-5
-10
-15
ppm

existence of a single alkaline species in Ala80cyt *c* demonstrates that the creation of a less hindered distal pocket can preferentially stabilize one of the alkaline conformations. These data suggest that the Lys coordinated in the alkaline form of Ala80cyt *c* is not Lys79, but rather one (most likely Lys73) whose coordination would require a significant conformational change, since no other Lys is near the heme pocket (Figure 4.1). It seems unlikely that Lys79, which does not coordinate to the Ala80cyt *c* heme at neutral pH, could compete with hydroxide as a ligand for iron as the pH is raised. It is more reasonable that hydroxide displacement by a Lys would occur if accompanied by a large conformational rearrangement.

Conclusions

Investigating the pH dependence of the UV/vis absorption spectra and of the hyperfine-shifted resonances of an axial-ligand (Met80Ala) variant of cyt *c* has provided valuable information on the pH dependence of the heme coordination of the mutants. The changes observed in the UV/vis absorption spectra of semisynthetic Ala80 and Phe67Ala80cyt *c* suggest that these proteins undergo a transition from low-spin to high-spin as pH is lowered, with midpoint transitions at pH 6.3 for both proteins. Monitoring the UV/vis spectrum of *S. c.* Fe(III)Ala80 as a function of pH indicates that a similar transition takes place with a slightly higher pK of 6.6. By investigating the pH dependence of the ^{1}H NMR spectra of *S. c.* Fe(III)Ala80cyt *c*, it has been confirmed that the protein undergoes a transition from low-spin to high-spin as the pH is lowered below 7, consistent with the result obtained from absorption spectroscopy. This result, however, contradicts the MCD data (Chapter 3) which suggest that *S. c.* Fe(III)Ala80 is high-spin at pH 7. At the lowest pH values examined, the spectra are similar for both *S. c.* WTcyt *c* and for *S. c.* Ala80cyt *c*, indicating that Met80 does not bind the heme iron at low pH in the wild-type protein. The short T_1 values for the heme methyl protons at pH 7.0 support

the hypothesis that hydroxide is a heme ligand at this pH. At pH 11, the spectrum of *S. c.* Ala80cyt *c* resembles that of one of the two alkaline forms of the wild-type protein.

The structure and folding of cyt *c* have long been subjects of investigation. The alkaline (pH > 9) forms of the protein have been studied for 50 years, but, remarkably, their structures are still a mystery. The preferential stabilization of one of the two alkaline isoforms by mutation of Met80 to Ala is consistent with the hypothesis that formation of one of them requires a significant conformational change. These findings provide an example of the influence that metal-ligand interactions can have on the conformational energetics of metalloproteins.

References and Notes

(1) The work in this chapter was done in collaboration with Dr. Paola Turano, Prof. Lucia Banci, and Prof. Ivano Bertini at the University of Florence.

(2) Theorell, H.; Åkesson, Å. *J. Am. Chem. Soc.* **1941,** *63,* 1804-1820.

(3) Dickerson, R. E.; Timkovich, R. In *The Enzymes*; Boyer, P., Ed.; Academic Press: New York, 1975; Vol. XI, pp 382-421.

(4) Moore, G. R.; Pettigrew, G. W. *Cytochromes c; Evolutionary, Structural and Physicochemical Aspects*; Springer-Verlag: Berlin, 1990.

(5) Goto, Y.; Hagihara, Y.; Hamada, D.; Hoshino, M.; Nishii, I. *Biochemistry* **1993,** *32,* 11878-11885.

(6) Gadsby, P. M. A.; Peterson, J.; Foote, N.; Greenwood, C.; Thomson, A. J. *Biochem. J.* **1987,** *246,* 43-54. Brautigan, D. L.; Feinberg, B. A.; Hoffman, B. M.; Margoliash, E.; Peisach, J.; Blumberg, W. E. *J. Biol. Chem.* **1977,** *252,* 574-582.

(7) Wallace, C. J. A. *Biochem. J.* **1984,** *217,* 601-604. Redfield, A. G.; Gupta, R. K. *Cold Spring Harbor Symp. Quant. Biol.* **1971,** *36,* 405-411. Stellwagen, E.; Babul, J.; Wilgus, H. *Biochim. Biophys. Acta* **1975,** *405,* 115-121.

(8) Hong, X. L.; Dixon, D. W. *FEBS Lett.* **1989,** *246,* 105-108.

(9) Ferrer, J. C.; Guillemette, J. G.; Bogumil, R.; Inglis, S. C.; Smith, M.; Mauk, A. G. *J. Am. Chem. Soc.* **1993,** *115,* 7507-7508.

(10) Mauk, A. G., personal communication.

(11) Berghuis, A. M.; Brayer, G. D. *J. Mol. Biol.* **1992,** *223,* 959-976.

(12) Antonini, E.; Brunori, M. *Hemoglobin and Myoglobin in Their Reactions with Ligands*; North-Holland: Amsterdam, 1971.

(13) Giacometti, G. M.; Ascenzi, P.; Brunori, M.; Rigatti, G.; Giacometti, G.; Bolognesi, M. *J. Mol. Biol.* **1981,** *151,* 315-319.

(14) Ikeda-Saito, M.; Hori, H.; Andersson, L. A.; Prince, R. C.; Pickering, I. J.; George, G. N.; Sanders, C. R., II; Lutz, R. S.; McKelvey, E. J.; Mattera, R. *J. Biol. Chem.* **1992,** *267,* 22843-22852.

(15) Bogumil, R.; Maurus, R.; Hildebrand, D. P.; Brayer, G. D.; Mauk, A. G. *Biochemistry* **1995,** *34,* 10483-10490.

(16) Morikis, D.; Champion, P. M.; Springer, B. A.; Egeberg, K. D.; Sligar, S. G. *J. Biol. Chem.* **1990,** *265,* 12143-12145.

(17) Giacometti, G. M.; Da Ros, A.; Antonini, E.; Brunori, M. *Biochemistry* **1975,** *14,* 1584-1588.

(18) Bertini, I.; Luchinat, C. *NMR of Paramagnetic Molecules in Biological Systems*; Benjamin/Cummings: Menlo Park, 1986.

(19) Bertini, I.; Turano, P.; Vila, A. J. *Chem. Rev.* **1993,** *93,* 2833-2932.

(20) McConnell, H. M.; Robertson, R. E. *J. Chem. Phys.* **1958,** *29,* 1361-1365. Kurland, R. J.; McGarvery, B. R. *J. Magn. Res.* **1970,** *2,* 286-304.

(21) McConnell, H. M.; Chestnut, D. B. *J. Chem. Phys.* **1958,** *28,* 107-117. McConnell, H. M. *J. Chem. Phys.* **1956,** *24,* 764-766.

(22) Horrocks, W. D., Jr.; Hall, D. D. *Inorg. Chem.* **1971,** *10,* 2368-2370. Vega, A. J.; Fiat, D. *Mol. Phys.* **1976,** *31,* 347-355.

(23) Koenig, S. H. *J. Magn. Reson.* **1982,** *47,* 441-453. Bloembergen, N. *J. Chem. Phys.* **1957,** *27,* 572-573.

(24) Inubushi, T.; Becker, E. D. *J. Magn. Reson.* **1983,** *51,* 128-133.

(25) Vold, R. L.; Waugh, J. S.; Klein, M. P.; Phelps, D. E. *J. Chem. Phys.* **1968,** *48,* 3831-3832.

(26) Banci, L.; Bertini, I.; Luchinat, C. *Nuclear and Electron Relaxation; The Magnetic Nucleus-Unpaired Electron Coupling in Solution*; VCH: Weinheim, 1991.

(27) Banci, L. In *Biological Magnetic Resonance*; Berliner, L. J., Reuben, J., Eds.; Plenum Press: New York, 1993, Vol. 12, pp 79-111.

(28) La Mar, G. N.; de Ropp, J. S. In *Biological Magnetic Resonance*; Berliner, L. J., Reuben, J., Eds.; Plenum Press: New York, 1993, Vol. 12, pp 1-78.

(29) Ferrer, J. C.; Turano, P.; Banci, L.; Bertini, I.; Morris, I. K.; Smith, K. M.; Smith, M.; Mauk, A. G. *Biochemistry* **1994,** *33,* 7819-7829. Vitello, L. B.; Erman, J. E.; Miller, M. A.; Mauro, J. M.; Kraut, J. *Biochemistry* **1992,** *31,* 11524-11535.

Satterlee, J. D.; Erman, J. E.; Mauro, J. M.; Kraut, J. *Biochemistry* **1990,** *29,* 8797-8804.

(30) La Mar, G. N.; Walker, F. A. In *The Porphyrins*; Dolphin, D., Ed.; Academic Press: New York, 1979; pp 61-157.

(31) Banci, L.; Bertini, I.; Eltis, L. D.; Pierattelli, R. *Biophys J.* **1993,** *65,* 806-813. Banci, L.; Bertini, I.; Marconi, S.; Pierattelli, R.; Sligar, S. G. *J. Am. Chem. Soc.* **1994,** *116,* 4866-4873.

(32) Banci, L.; Bertini, I.; Turano, P.; Tien, M.; Kirk, T. K. *Proc. Natl. Acad. Sci. U.S.A.* **1991,** *88,* 6956-6960. Banci, L.; Bertini, I.; Turano, P.; Ferrer, J. C.; Mauk, A. G. *Inorg. Chem.* **1991,** *30,* 4510-4516. Banci, L.; Bertini, I.; Pease, E. A.; Tien, M.; Turano, P. *Biochemistry* **1992,** *31,* 10009-10017.

(33) Feng, Y.; Roder, H.; Englander, S. W. *Biophys. J.* **1990,** *57,* 15-22. Gao, Y.; Boyd, J.; Williams, R. J. P.; Pielak, G. J. *Biochemistry* **1990,** *29,* 6994-7003.

(34) Pearce, L. L.; Gartner, A. L.; Smith, M.; Mauk, A. G. *Biochemistry* **1989,** *28,* 3152-3156.

Chapter 5

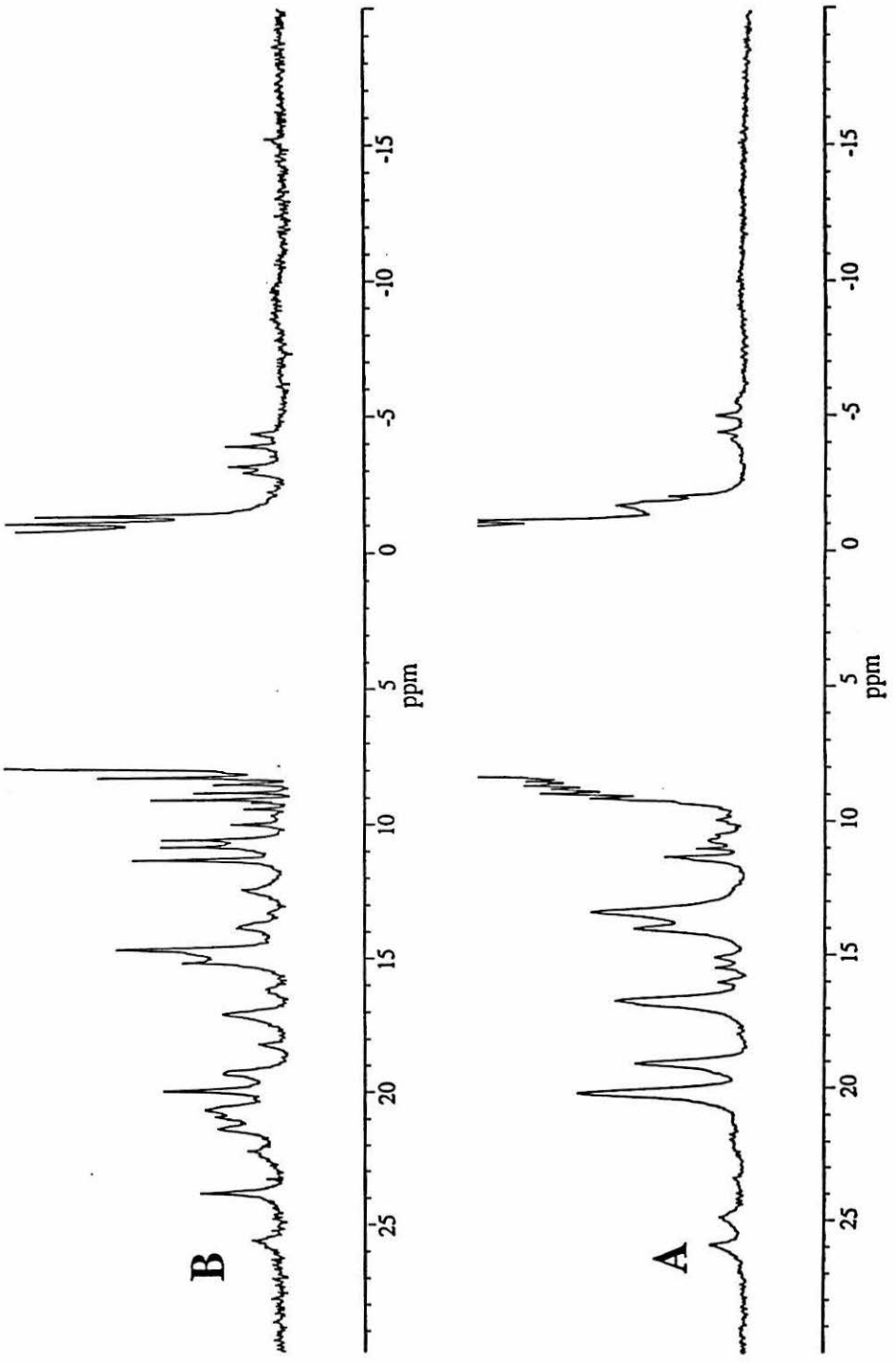
Paramagnetic ^{1}H NMR Spectroscopy of the Cyanide Derivative of *Saccharomyces cerevisiae* Ala80cytochrome c[1]

Background

The results presented in Chapter 4 of this thesis illustrate the utility of ^{1}H NMR spectroscopy for probing the spin state and the conformational equilibria of oxidized cyts *c* through a wide range of pH values. In the research discussed in the final two chapters of this thesis, ^{1}H NMR is used to gain more detailed information on the electronic and molecular structure of *S. c.* Ala80cyt *c*. As is evident upon inspection of the data presented in Chapter 4 and in Figure 5.1, the linewidths and complexity of the ^{1}H NMR spectra of Fe(III)Ala80cyt *c* when the sixth iron coordination site is occupied by OH^- or H_2O are not conducive to a detailed NMR study, in particular to the assignment of the hyperfine-shifted signals. The ^{1}H NMR spectrum of Fe(III)Ala80 below pH 7 indicates a high-spin species; the signals due to protons that are coupled with the metal ion are dramatically shifted and broadened (Figure 4.7B–F). As with the analogous Mb and peroxidase systems,[2] the detailed assignment is very difficult, if not impossible, unless specific nuclear labeling is used. The protons of the low-spin species above pH 7 have short nuclear T_1 values, similar to those of the high-spin species of metMb, and provide relatively broad resonances. This species also exhibits slow exchange both with the low-pH species and with an additional low-spin species on the high-pH side, complicating the spectra considerably.

In order to increase nuclear T_1 values, giving relatively narrow hyperfine-shifted lines, and to minimize the complication of ionization and conformational equilibria, the cyanide derivative of *S. c.* Fe(III)Ala80cyt *c* has been chosen as the subject of the NMR investigation presented here and in Chapter 6. A number of groups have found that the cyanide complexes of heme protein mutants exhibit favorable nuclear T_1 values and fewer ionization and conformational equilibria than the ferric protein without added cyanide.[2,3,8-15] The cyanide derivative has the additional advantage that the heme-bound cyanide can interact with hydrogen-bond donors in the heme pocket, making it an acceptable model for a heme-O_2 unit. This allows us to probe ligand-heme cavity

Figure 5.1 600 MHz ^{1}H NMR spectra of *S. c.* Fe(III)Ala80cyt *c* recorded at (A) 303 K and (B) 306 K. The 1-mM sample was in 50 mM NaP_i (pH 7.0). The broad hyperfine-shifted peaks between 10 and 25 ppm are characteristic of a low-spin Fe(III) protein with a weak sixth heme ligand, and the significant differences seen between the spectra at the two temperatures indicate the presence of conformational equilibria.


B
A
25
20
15
10
5
0
-5
-10
-15
ppm

interactions with the goal of elucidating the mechanisms of stabilization of the dioxygen adduct of Fe(II)Ala80cyt *c*.

Materials and Methods

^{1}H NMR Spectroscopy

Wild-type cyt *c* (Ser102-iso-1-cyt *c*) and Ala80cyt *c* (Gln39Ala80Ser102-iso-1-cyt *c*) were expressed in *Saccharomyces cerevisiae* and purified as described in Chapter 3. ^{1}H NMR sample preparation was described in Chapter 4. For the experiments in this chapter, samples from 1–3 mM were used. Nonselective T_1 experiments were performed as described in Chapter 4. The cyanide derivative of Fe(III)Ala80cyt *c* was prepared by adding a small (~ 1.1x) excess of KCN directly to the oxidized protein sample in the NMR tube. The cyanide derivative of Fe(III)WTcyt *c* was prepared by adding a larger (~ 5x) excess of KCN to the sample.

^{1}H NMR spectra were recorded on a Bruker MSL 200 (operating at 200.13 MHz) or a Bruker AMX 600 (operating at 600.13 MHz) spectrometer. The 200 MHz 1D ^{1}H NMR spectra were recorded using a superWEFT[16] pulse sequence with a recycle delay of 200–250 ms. The 600 MHz 1D ^{1}H NMR spectra were obtained using presaturation during the relaxation delay to eliminate the water signal.

The ^{1}H nuclear Overhauser effect (NOE) experiments (at 200 and 600 MHz) were performed with the superWEFT pulse sequence for water signal suppression and were collected using the published procedure.[17] Difference spectra were collected by applying the decoupler frequency on- and off- resonance according to the scheme: ω, $\omega + \delta$, ω, $\omega - \delta$, where ω is the frequency of the irradiated signal and δ is the offset for the off-resonance irradiation. The application of the off-resonance decoupling symmetrically on both sides of the saturated signal minimizes the off-resonance effect of the saturation. The phase of the receiver was alternated accordingly so that a difference free induction decay (FID) was collected directly.

TPPI NOESY[18] spectra were recorded with presaturation of the solvent signal during both the relaxation delay and the mixing time. To optimize the detection of cross peaks involving fast-relaxing nuclei, NOESY maps at 295 and 303 K in D_2O (90% D_2O/10% H_2O) were recorded on the full spectral width (59.5 ppm) with a recycle time of 550 ms and a mixing time of 15 ms. To optimize the detection of connectivities in the diamagnetic region, NOESY maps in D_2O and H_2O (90% H_2O/10% D_2O) were recorded on a smaller spectral width (30 ppm) with a recycle time of 800 ms and mixing times of 25 and 100 ms. Analogously, TPPI TOCSY[19] experiments with presaturation during the relaxation delay were recorded on a spectral width of 59.5 ppm (recycle time of 800 ms and spin-lock time of 18 ms) and 30 ppm (recycle time of 800 ms and spin-lock times of 30, 60, and 90 ms). A magnitude COSY[20] (MCOSY) map was recorded in H_2O solution on a 30 ppm spectral width (recycle time of 800 ms). A NOE-NOESY experiment with a mixing time of 70 ms was performed by using the conventional NOESY sequence preceded by (180° pulse)-(τ delay) to select relatively fast-relaxing resonances. During the τ delay (200 ms) the 5-CH_3 signal was irradiated selectively. Analogous to the 1D NOE experiments described above, a difference map was obtained by applying the decoupler frequency on- and off-resonance, alternately, during the τ delay. The phase of the receiver was alternated accordingly, so that difference FIDs were directly collected.

All 2D spectra consisted of 4096 data points in the F2 dimension. From 800 to 1024 experiments were recorded in the F1 dimension, using 64–192 scans per experiment. Raw data were multiplied in both dimensions by a pure cosine-squared (NOESY, TOCSY, NOE-NOESY) or a pure sine-squared (COSY) bell window function and Fourier-transformed to obtain 2048 x 2048 real data points. A polynomial baseline correction was applied in both directions. Chemical shifts were calibrated to DSS through the residual water signal. Data processing was performed by using the standard Bruker software package. The 2D maps were analyzed on IBM RISC 6000 computers with the program XEASY.[21]

EPR Spectroscopy

The EPR spectra of the CN-Fe(III)WTcyt *c* and CN-Fe(III)Ala80cyt *c* (50% 100 mM NaP_i, pH 7, 50% glycerol) were measured at 6 K with a Bruker ER200 spectrometer operating at 9.6 GHz.

Results

Assignment of Heme Substituents

S. c. CN-Fe(III)Ala80cyt c

The 1D ^1H NMR spectrum of *S. c.* CN-Fe(III)Ala80cyt *c* is shown in Figure 5.2A. The spectrum has considerably narrower hyperfine-shifted lines than those of Fe(III)Ala80 without added cyanide (Figure 5.1). In addition, it is clear that one major species of CN-Ala80 exists. The assignment of the heme substituents began with the four heme methyl signals in the ^1H NMR spectrum of CN-Fe(III)Ala80cyt *c* at 303 K, which are the well-resolved signals of intensity 3 at 22.5, 19.5, 15.4, and 11.3 ppm (Figure 5.2A). The resonances at 22.5 and 15.4 both have strong dipolar connectivities with a resonance at -4.3 ppm (observed in a NOESY map, Figure 5.3, and checked through a 1D NOE experiment, Figure 5.4). Inspection of the *S. c.* WTcyt *c* X-ray crystal structure[22] reveals that the only proton that can be within NOE distance of two heme methyls is the δ-meso proton that is equidistant from 1-CH_3 and 8-CH_3. The NOESY map shown in Figure 5.3 was recorded with a mixing time of 15 ms to enhance the detection of cross peaks between protons with short T_1 values. A NOESY spectrum taken with a long (100-ms) mixing time (Figure 5.5) reveals dipolar connectivities between the resonance at 15.4 ppm and a number of amino acid side chains, revealed by inspection of TOCSY maps to be four different Leu residues and the ring of a Phe. By inspection of the X-ray crystal structure of WTcyt *c*, the methyl signal at 15.4 ppm can be assigned to 1-CH_3, which is close to the side chains of Leu32, Leu68, Leu94, Leu98, and Phe10. The assignment of

Figure 5.2 600 MHz ^{1}H NMR spectra (303 K, 50 mM NaP_i (uncorrected pH 7.0), D_2O) of (A) *S. c.* CN-Fe(III)Ala80cyt *c* and (B) *S. c.* CN-Fe(III)WTcyt *c*. The inset of (A) shows the spectrum of *S. c.* CN-Ala80 recorded in H_2O. The arrow indicates the Tyr67 hydroxyl proton resonance.

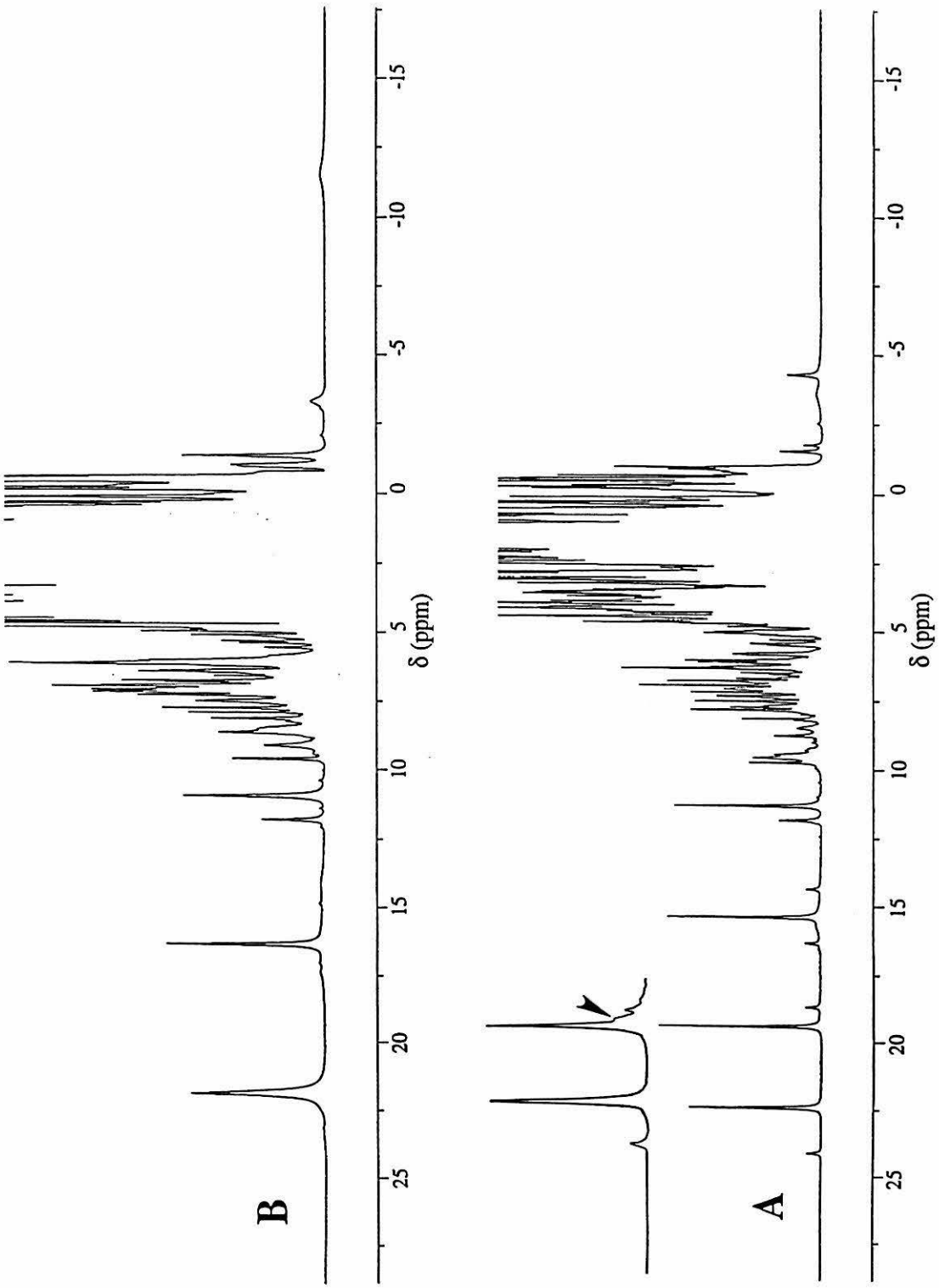

B
A
δ (ppm)
δ (ppm)
25
20
15
10
5
0
-5
-10
-15

Figure 5.3 600 MHz ^{1}H NMR TPPI NOESY map of CN-Fe(III)Ala80cyt *c* (303 K, 50 mM NaP_i (uncorrected pH 7.0), D_2O) recorded over a spectra width of 59.5 ppm using water presaturation during the relaxation delay (550 ms) and the mixing time (15 ms). Resolved cross peaks between heme resonances are labeled. The short mixing time optimizes detection of cross-peaks between strongly-relaxed nuclei.

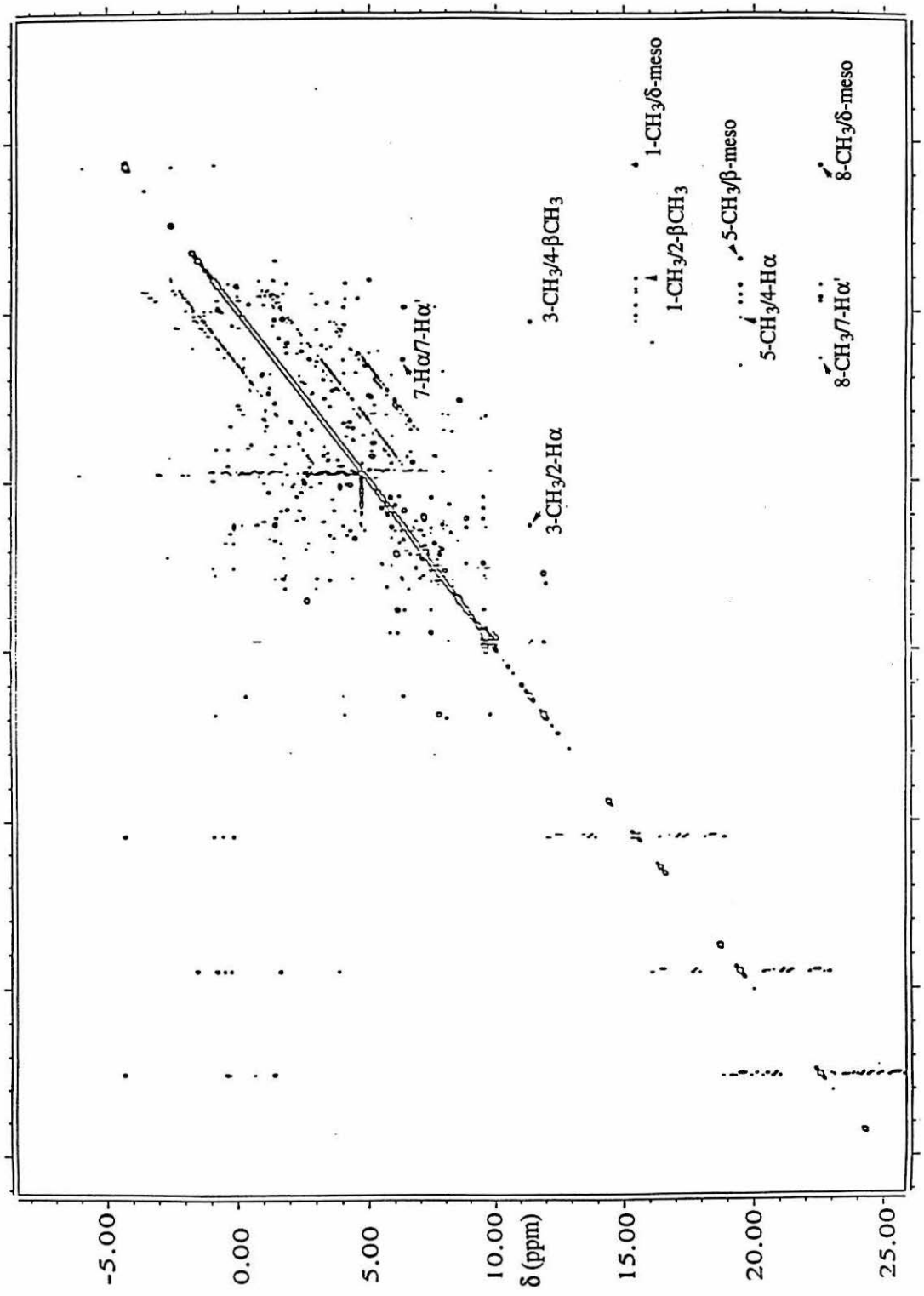

7-Hα/7-Hα'
3-CH3/4-βCH3
3-CH3/2-Hα
1-CH3/δ-meso
1-CH3/2-βCH3
5-CH3/β-meso
5-CH3/4-Hα
8-CH3/7-Hα'
8-CH3/δ-meso
δ (ppm)
δ (ppm)

Figure 5.4 (A) 1D NOE difference spectrum of *S. c.* CN-Fe(III)Ala80cyt *c* (303 K, 50 mM NaP_i (uncorrected pH 7.0), D_2O) obtained upon saturation of the δ-meso proton resonance (-4.3 ppm). The spectrum was recorded by using the superWEFT pulse sequence with a recycle delay of 500 ms and a τ of 500 ms. (B) 600 MHz ^1H NMR reference spectrum.

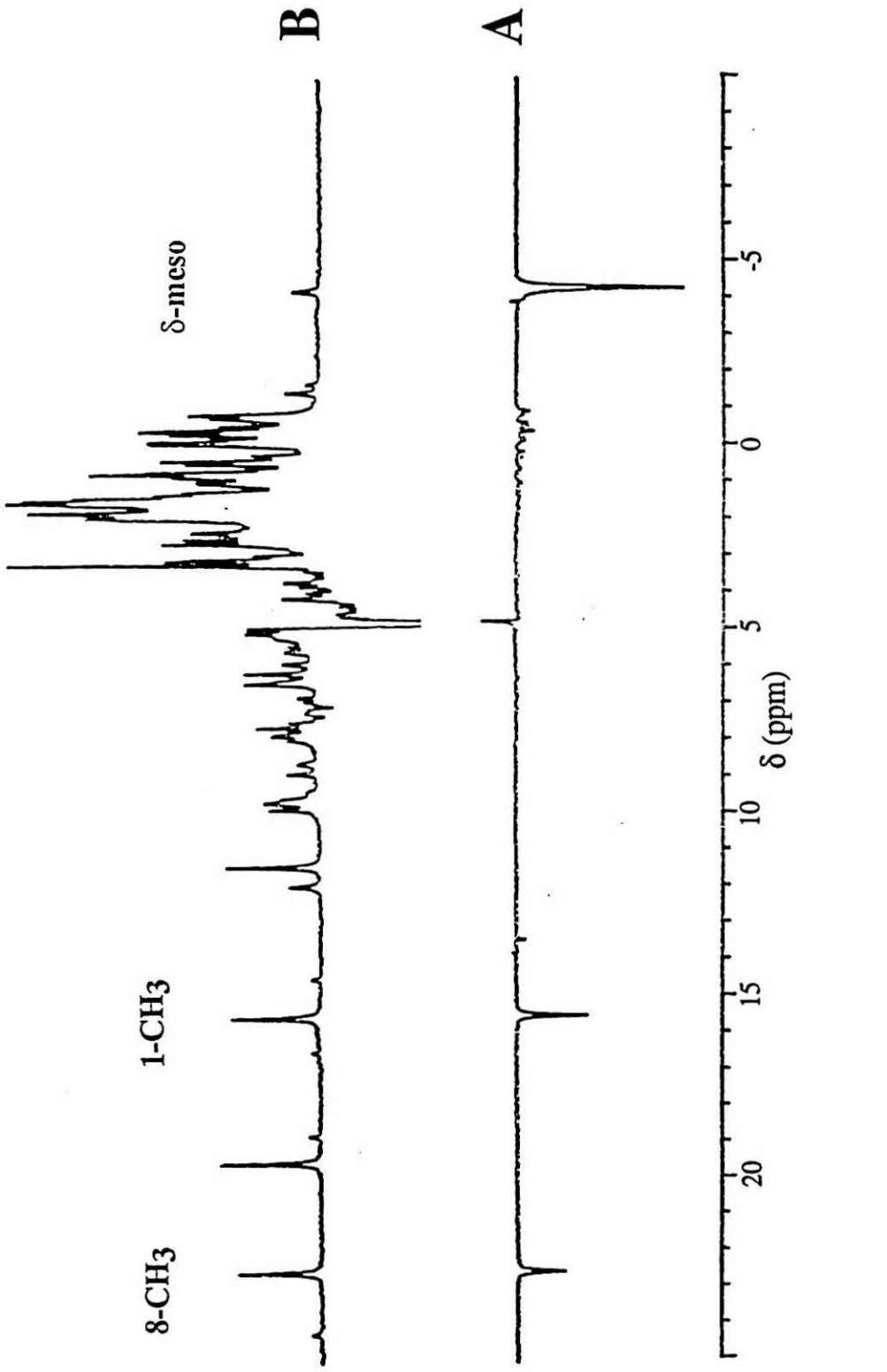
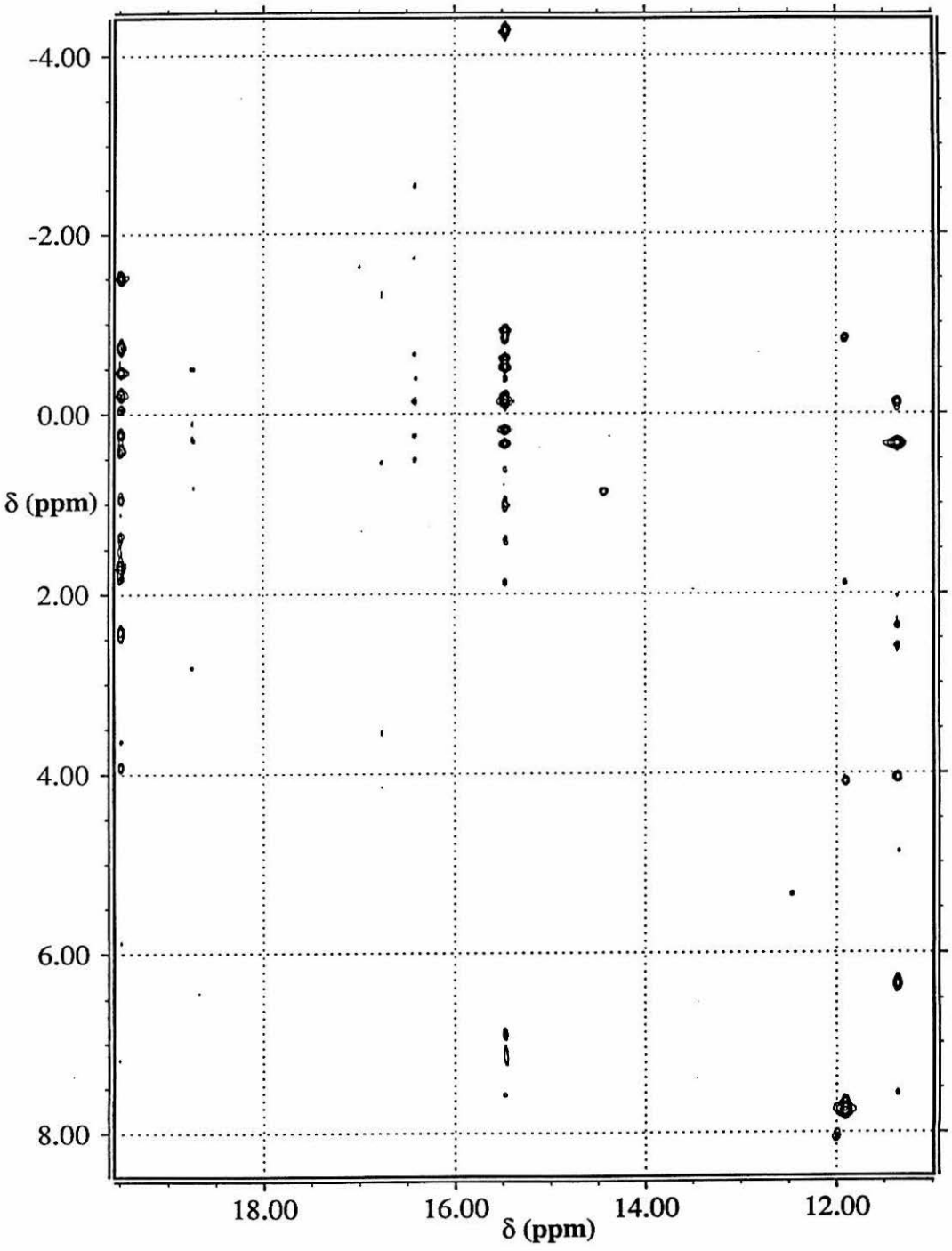

B
δ-meso
1-CH3
8-CH3
A
20
15
10
5
0
-5
δ (ppm)

Figure 5.5 600 MHz ^{1}H NMR TPPI NOESY spectrum of *S. c.* CN-Fe(III)Ala80cyt *c* (303 K, 50 mM NaP_i (uncorrected pH 7.0), D_2O) recorded with a long mixing time (100 ms). Cross-peaks in this region are due to dipolar connectivities between the four heme methyl resonances and nearby amino acid residues, as well as other heme substituents. See text for details. The 8-CH_3 resonance is folded in to the spectrum.


-4.00
-2.00
0.00
δ (ppm)
2.00
4.00
6.00
8.00
18.00
16.00
14.00
12.00
δ (ppm)

the resonance at 22.5 ppm to 8-CH_3 is consistent with its NOESY connectivities with Ile, Trp, and Met side chains (Ile53, Trp59, and Met64 from the crystal structure of WTcyt *c*).

The two remaining hyperfine-shifted resonances of intensity 3 at 19.5 and 11.3 ppm are attributable to the heme methyls 3 and 5. It is possible to distinguish between them on the on the basis of their dipolar connectivities with other protein residues, expected on the basis of the crystal structure of WTcyt *c*. The resonance at 19.5 ppm is assigned to the 5-CH_3 protons, since it has NOESY connectivities with a Gly (Gly29) and with some of the side-chain protons of a Lys (Lys79). A NOE-NOESY experiment (Figure 5.6) obtained by selective saturation of the resonance at 19.5 ppm was particularly helpful in identifying the connectivities with the Lys79 side chain; several of the protons coupled with the irradiated methyl (Figure 5.6, bottom inset) are also coupled with each other. Some of the latter connectivities are also detected in TOCSY experiments, confirming that these signals are due to protons in the same residue, confirmed to be Lys79 by the sequence-specific assignment discussed in detail in Chapter 6. The assignment of the resonance at 11.3 ppm to 3-CH_3 is validated by dipolar connectivities with the ring and β-CH_2 protons of a Phe (Phe82).

The remaining heme assignments were made by analyzing NOESY connectivities among heme substituents starting from positions 1 and 3 and continuing around the heme periphery as illustrated in Figure 5.7. The assignment of the 6-propionate required the use of the NOE-NOESY experiment (Figure 5.6), which allowed the identification of the propionate chain Hα, Hα', Hβ resonances. The 6-Hβ' resonance remains unassigned. The complete assignment of the 7-propionate was obtained from the NOESY connectivities observed from 8-CH_3 to the Hα, Hα', Hβ' of 7-propionate and between the 7-propionate resonances (Hα/Hα', Hα/Hβ, Hα/Hβ', Hα'/Hβ, Hα'/Hβ'). The Hβ/Hβ' connectivity was not detected due to the nearly degenerate shifts (0.6 and 0.8 ppm) of these two resonances.

Figure 5.6 600 MHz ^{1}H NMR TPPI NOE-NOESY map of *S. c.* CN-Fe(III)Ala80cyt *c* (303 K, 50 mM NaP_i (uncorrected pH 7.0), D_2O) recorded by using the following pulse sequence: $180°-\tau(200\ ms)-90°-t_1-90°-\tau_m(70\ ms)-90°-t_2$, with selective saturation of the 5-CH_3 resonance during the 200 ms τ delay. The signals that are directly coupled with the saturated one lie on the diagonal and at the $\omega 1$ frequency of the saturated signal (inset at bottom). The peaks in the inset are between the 5-CH_3 and (a) Gly29Hα1, (b) Lys79Hβ, Hβ', (c) heme 6-Hα, (d) heme 6-Hα', (e) Lys79Hδ1, (f) Val28Hγ2, (g) Lys79Hε2, (h) heme 6-Hβ, (i) heme 4-Hα, (j) Lys79Hγ2, (k) Val28Hγ1, (l) heme β-meso, (m) Ala80Hα. Off the diagonal are the cross peaks between the diagonal resonances and other protons close to them. Dotted lines indicate peaks due to the Lys79 residue.

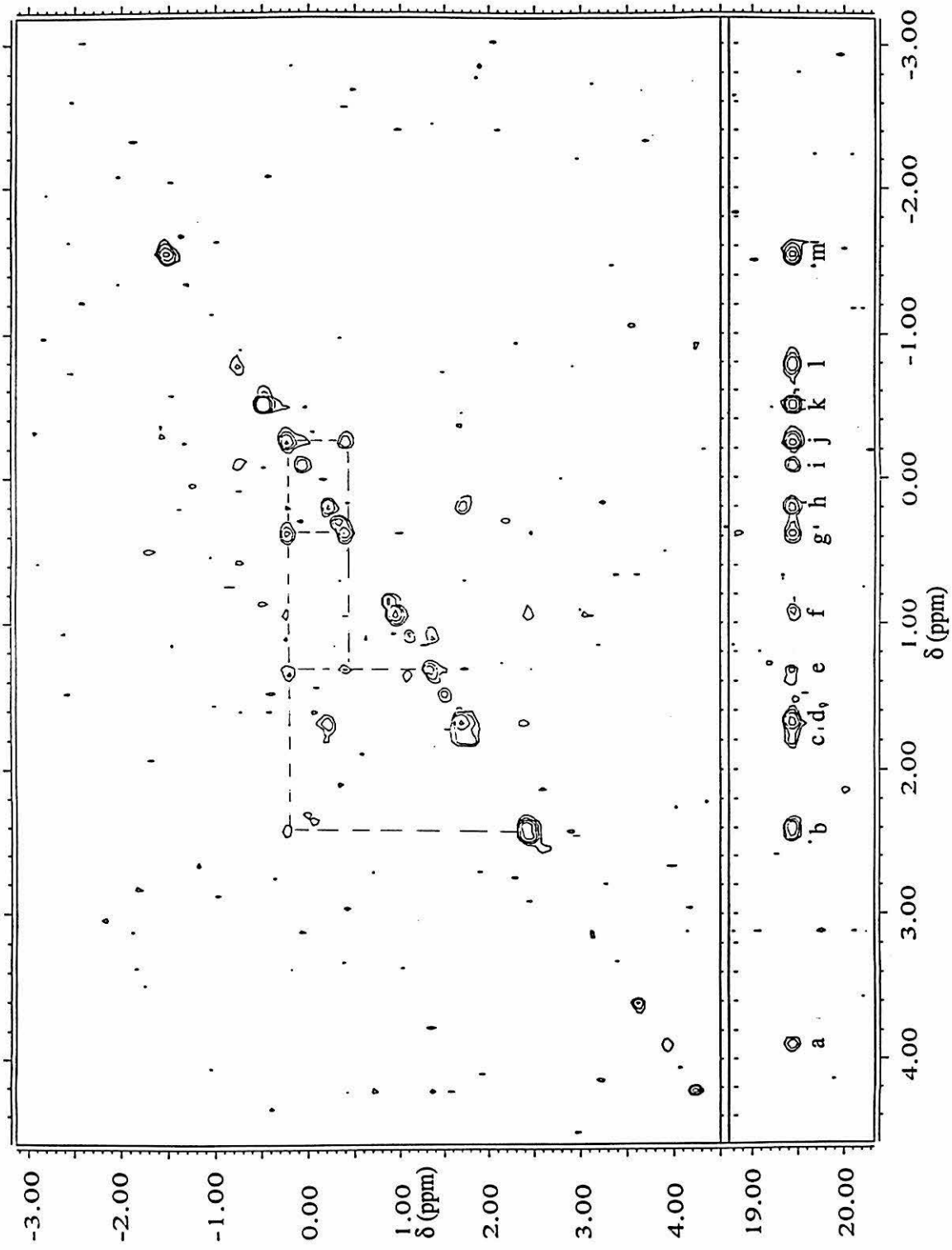

-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
19.00
20.00
δ (ppm)
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
δ (ppm)
a
b
c, d
e
f
g' h
i
j
k
l
m

Figure 5.7 Schematic representation of a *c*-type heme showing the dipolar connectivities observed among the heme resonances of *S. c.* CN-Fe(III)Ala80cyt *c*.

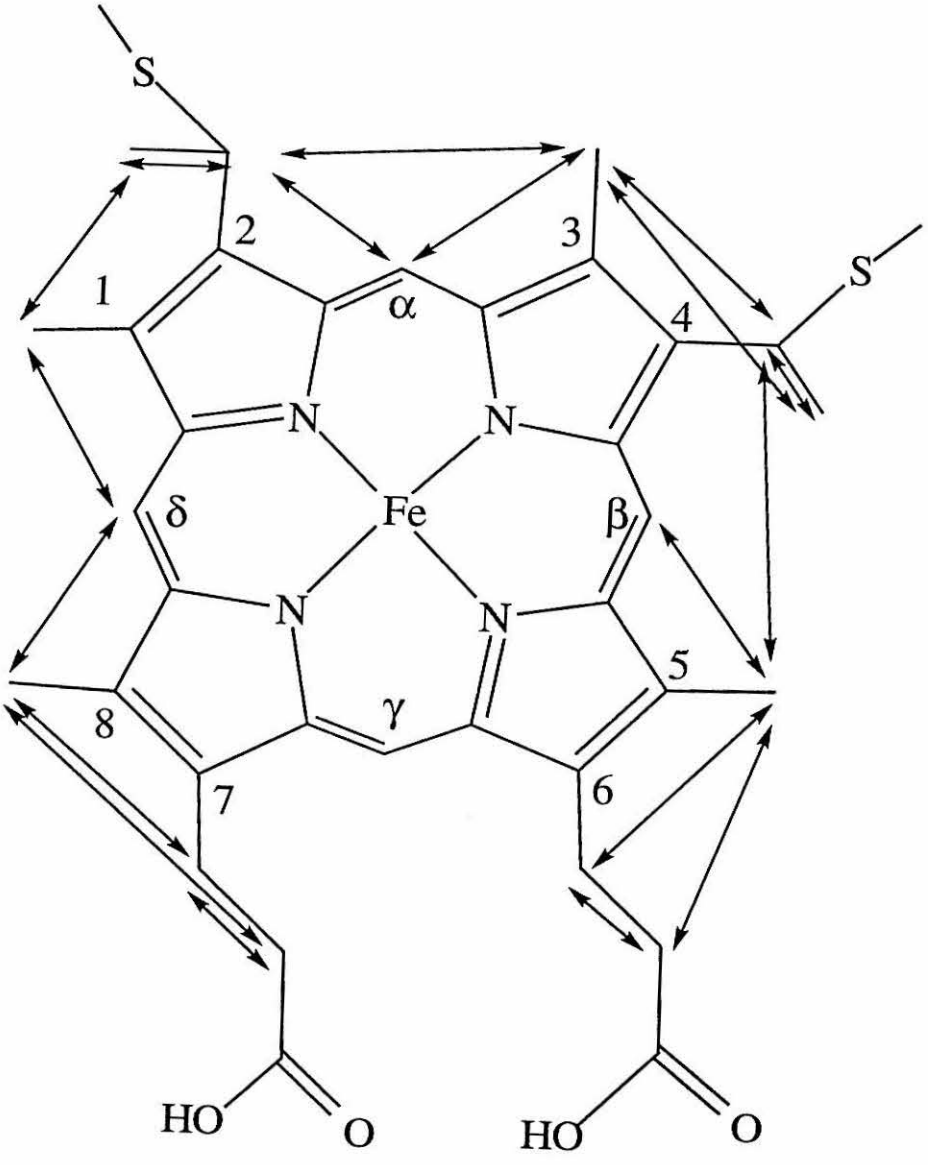

S
S
1
2
3
4
5
6
7
8
α
β
γ
δ
N
N
N
N
Fe
HO
O
HO
O

Table 5.1

Assignments for the ^1H NMR signals of the heme substituents of paramagnetic cyts c[a]

Assignment	WTcyt *c*, ppm	CN-WTcyt *c*, ppm	CN-Ala80cyt *c*, ppm
8–CH_3	34.2	22.2	22.5
5–CH_3	10.9	21.8	19.5
1–CH_3	8.0	16.5	15.4
3–CH_3	31.0	11.1	11.3
2–Hα	-0.7	*	6.3
2–CH_3	-2.1	*	-0.2
4–Hα	2.9	*	-0.1
4–CH_3	2.1	*	0.3
7–Hα	15.9	*	6.3
7–Hα'	12.8	*	1.4
7–Hβ	-0.2	*	0.6
7–Hβ'	1.4	*	0.8
6–Hα	2.9	*	1.8
6–Hα'	-1.4	*	1.6
6–Hβ	1.0	*	**
6–Hβ'	2.7	*	0.2
α–meso	2.6	*	4.1
β–meso	1.8	*	-0.8
γ–meso	8.1	*	**
δ–meso	2.3	-3.0	-4.3

[a]T = 303 K. All cyts *c* are from *Saccharomyces cerevisiae.*

*assignment not attempted.

**assignment attempted but not made.

S. c. CN-Fe(III)WTcyt c

A NOESY spectrum of *S. c.* CN-Fe(III)WTcyt *c* recorded at 600 MHz with a mixing time of 15 ms allows the assignment of the heme methyl signals, as reported in Table 5.1. The NOESY patterns from each of the heme methyl resonances are very

similar to those observed for *S. c.* CN-Fe(III)Ala80cyt *c* acquired under the same experimental conditions.

S. c. Fe(III)WTcyt c

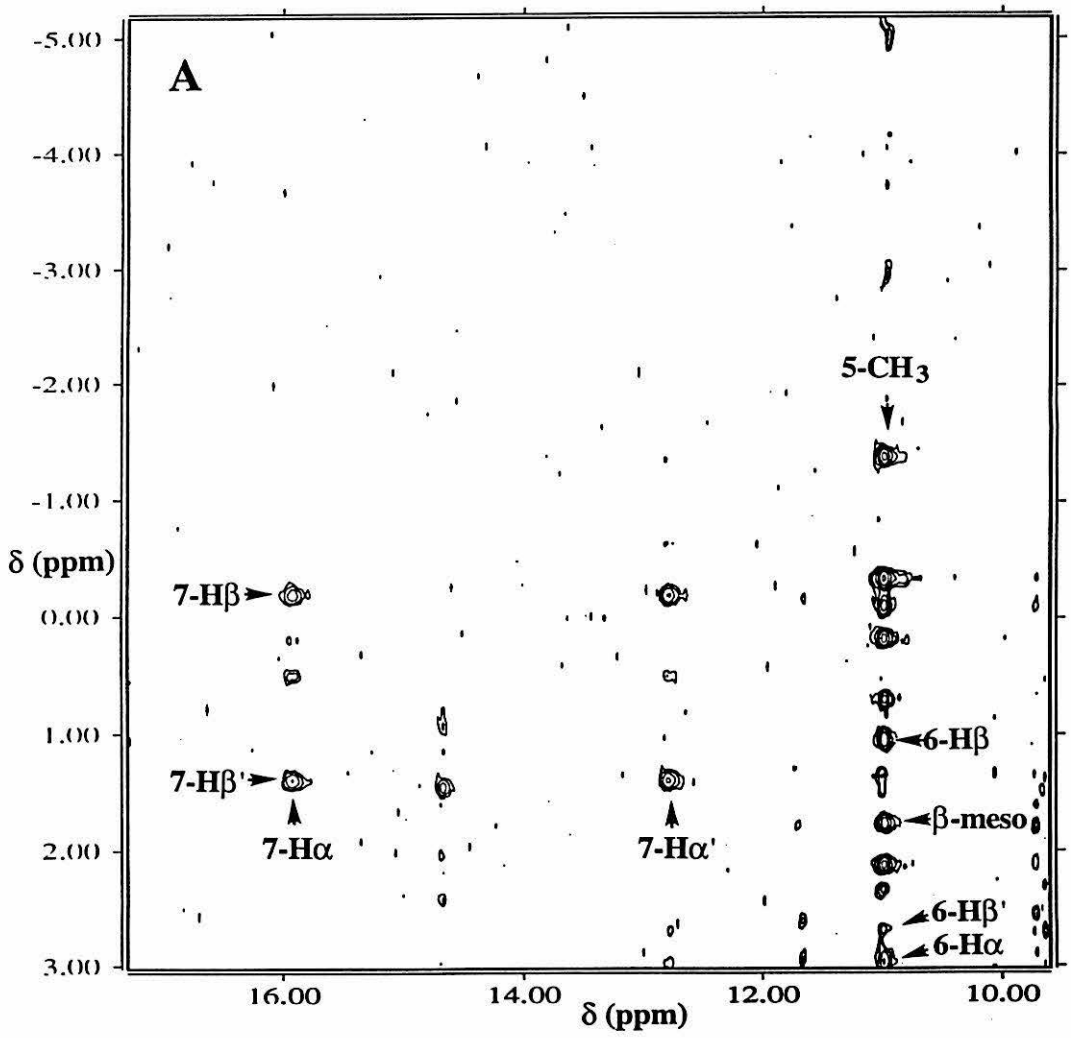
An extensive assignment of the proton resonances of *S. c.* Thr102cyt *c* is available in the literature.[23] This and other similar reports, however, do not include a complete assignment of the cyt *c* heme protons in the oxidized state. The partial assignment available in the literature and the methodology used in successfully obtaining a nearly complete assignment of the heme proton resonances of *S. c.* CN-Fe(III)Ala80 allowed the facile assignment of all of the heme resonances for Fe(III)WTcyt *c*, reported in Table 5.1. The assignments were made in a similar fashion as those reported above for CN-Ala80. In addition, the X-ray crystal structure was used to predict dipolar connectivities expected between heme substituents and nearby amino acids. Examples of these dipolar connectivities between heme substituents as observed in NOESY maps are shown in Figure 5.8 and 5.9. The assignment of the 6-propionate protons, which has not been reported before, was aided by the observation of a TOCSY pattern among the four protons in a spectrum taken with a short (30 ms) mixing time (Figure 5.9B).

Assignment of Tyr67

S. c. CN-Fe(III)Ala80cyt c

There is a broad, exchangeable signal at 19.3 ppm, visible in the 1D spectrum of Ala80cyt *c* in H_2O (inset of Figure 5.2). Upon saturation of this signal, NOEs on resonances at 7.8, 7.6, 7.3, and 7.2 ppm are observed. A TOCSY spectrum taken with a spin-lock time of 30 ms shows that these signals belong to a Tyr ring that is not flipping, or that is flipping slowly on the NMR timescale (Figure 5.10). Saturation of the Hε1 of this Tyr induces a NOE on the heme 7-propionate Hα' resonance. Other NOEs from this exchangeable signal are observed on resonances at 6.1, 8.8, and 9.5 ppm, which were

Figure 5.8 600 MHz ^{1}H NMR TPPI NOESY spectra of *S. c.* Fe(III)WTcyt *c* (303 K) in 50 mM NaP_i (pH 7.0) acquired with a 100 ms mixing time. Dipolar connectivities between heme substituents are indicated.


A
-5.00
-4.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
δ (ppm)
16.00
14.00
12.00
10.00
δ (ppm)
5-CH3
7-Hβ
7-Hβ'
7-Hα
7-Hα'
6-Hβ
β-meso
6-Hβ'
6-Hα

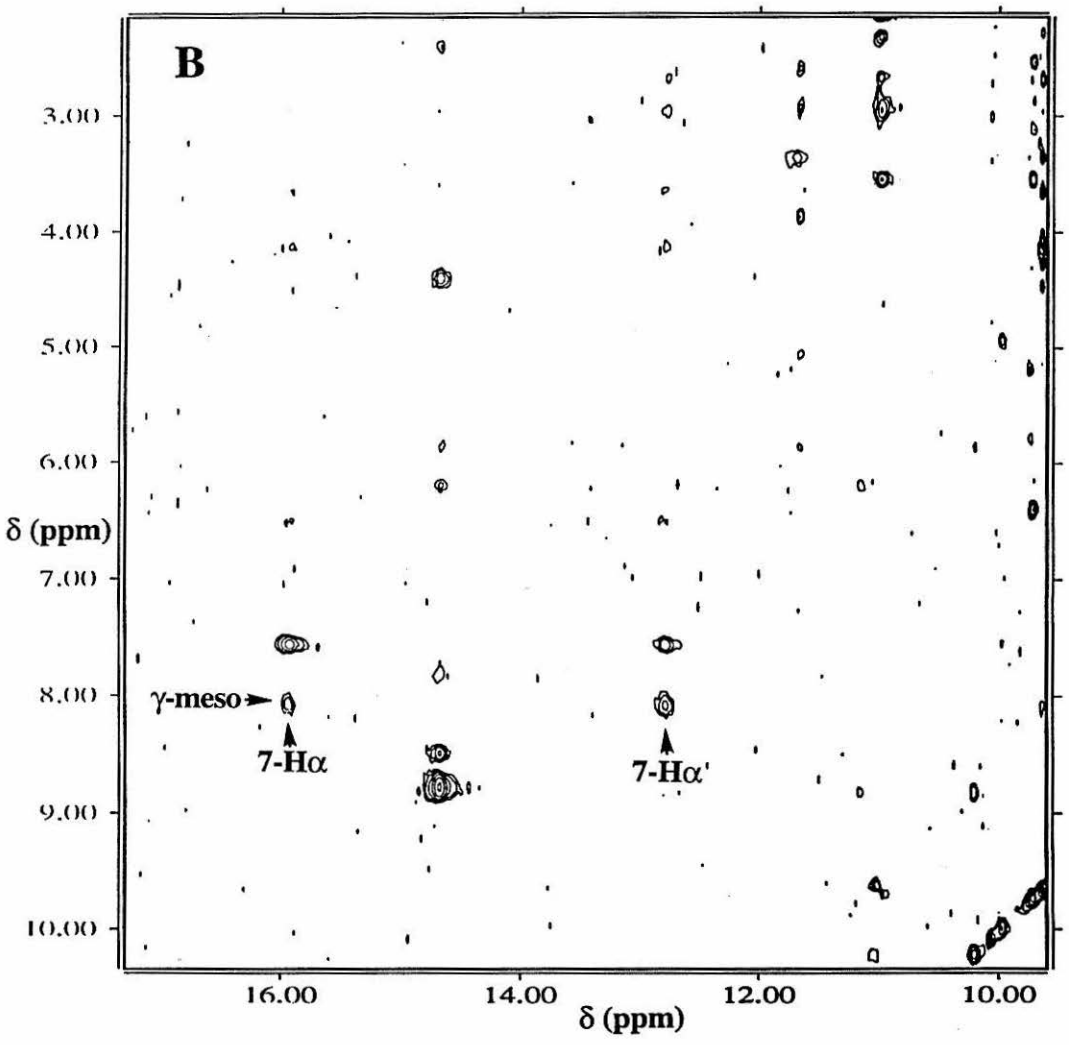
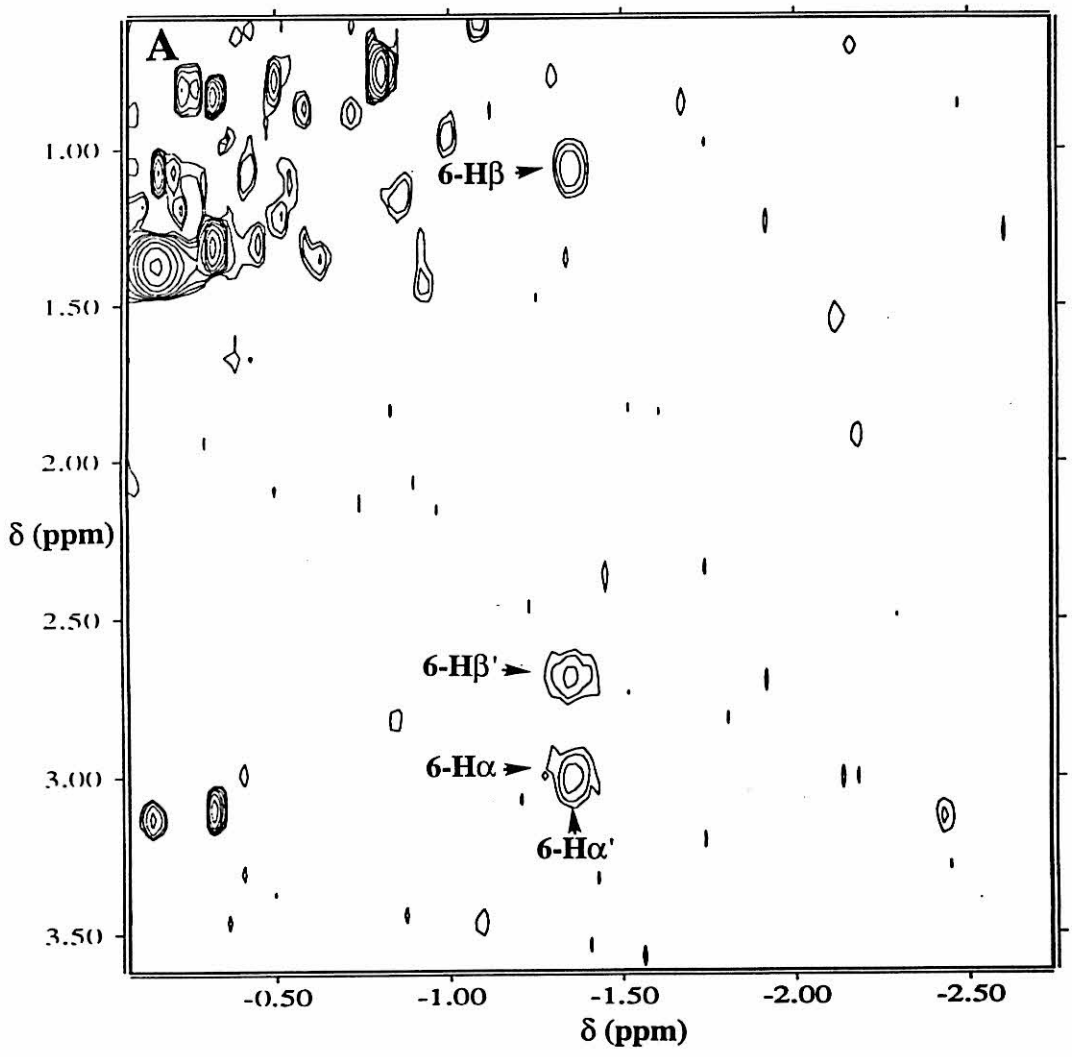

B
3.00
4.00
5.00
6.00
δ (ppm)
7.00
8.00
9.00
10.00
γ-meso
7-Hα
7-Hα'
16.00
14.00
12.00
10.00
δ (ppm)

Figure 5.9 (A) 600 MHz ^{1}H NMR TPPI TOCSY map of *S. c.* Fe(III)WTcyt *c* (303 K) in 50 mM NaP_i (pH 7.0) acquired with a spin-lock time of 30 ms. The spin pattern of the heme 6-propionate is indicated. (B) 600 MHz ^{1}H NMR TPPI NOESY map of *S. c.* Fe(III)WTcyt *c* (303 K) in 50 mM NaP_i (pH 7.0) acquired with a mixing time of 100 ms. Dipolar connectivities involving the heme 6-propionate and other heme protons are indicated.


A
6-Hβ
6-Hβ'
6-Hα
6-Hα'
1.00
1.50
2.00
2.50
3.00
3.50
δ (ppm)
-0.50
-1.00
-1.50
-2.00
-2.50
δ (ppm)

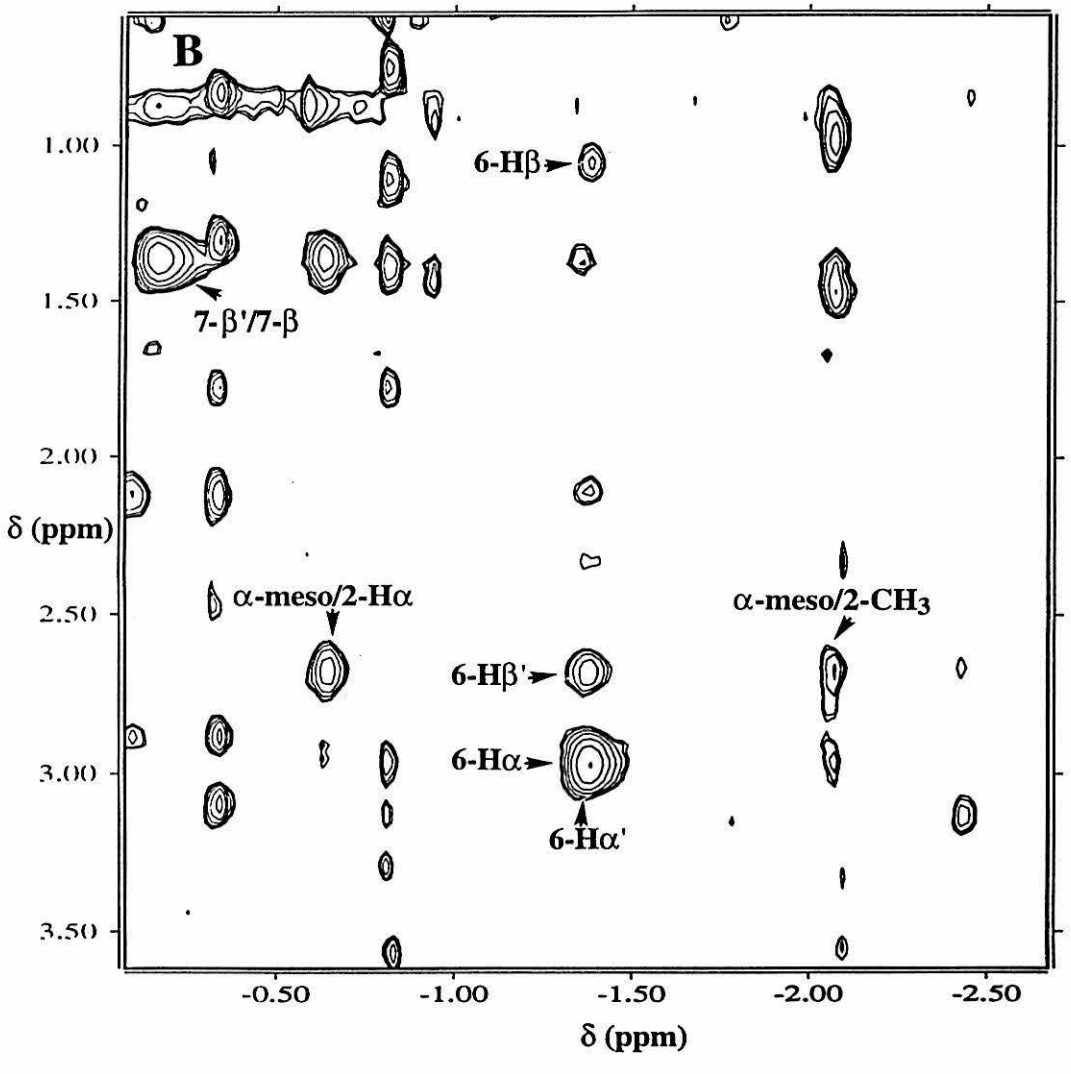
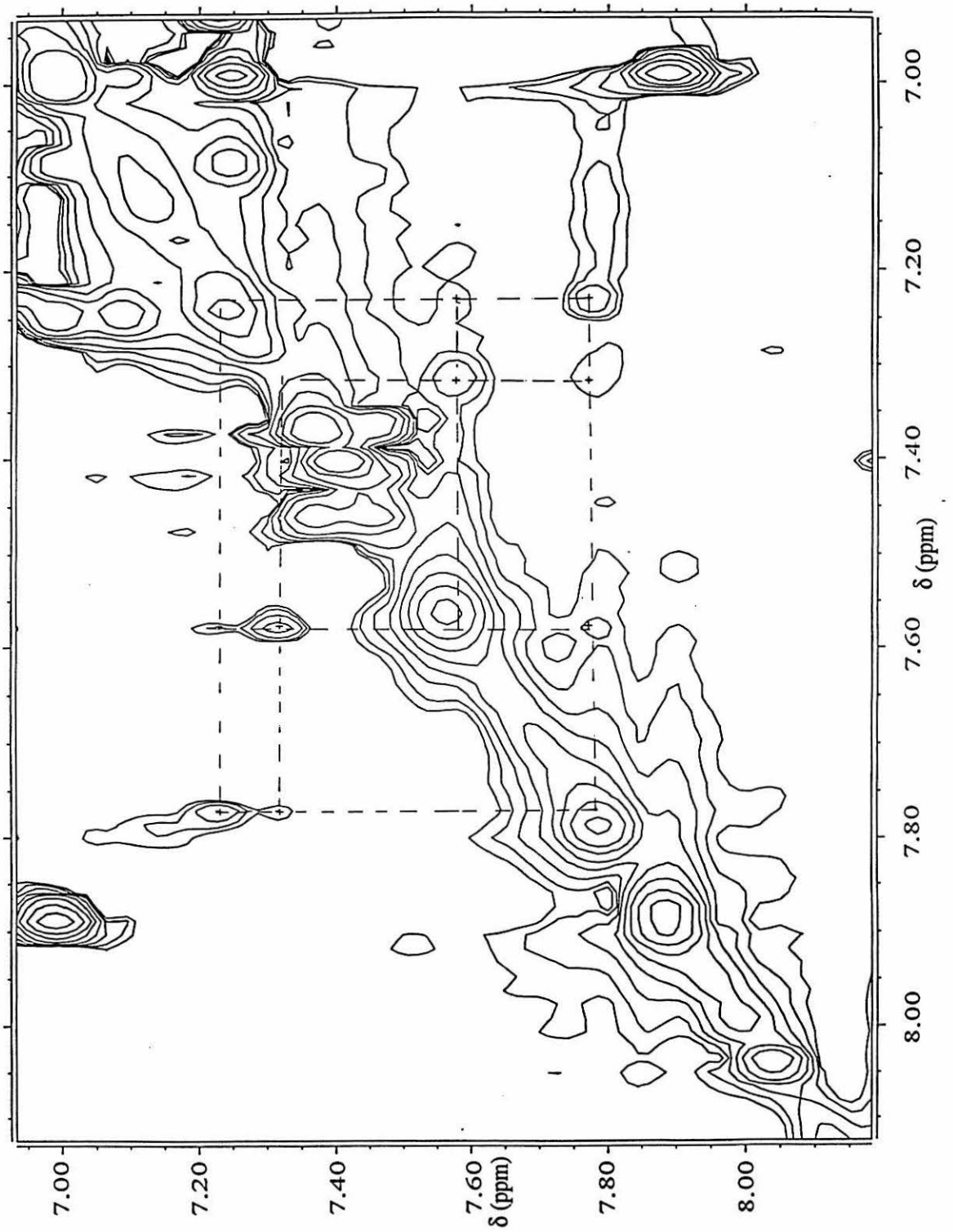

B
1.00
1.50
2.00
2.50
3.00
3.50
δ (ppm)
6-Hβ
7-β'/7-β
α-meso/2-Hα
α-meso/2-CH3
6-Hβ'
6-Hα
6-Hα'
-0.50
-1.00
-1.50
-2.00
-2.50
δ (ppm)

Figure 5.10 600 MHz ^{1}H NMR TPPI TOCSY map of *S. c.* CN-Fe(III)Ala80cyt *c* (303 K, 50 mM NaP_i (uncorrected pH 7.0), D_2O) recorded with a spin-lock time of 30 ms. Dotted lines define the spin pattern for the Tyr67 ring.


7.00
7.20
7.40
7.60
7.80
8.00
δ (ppm)
8.00
7.80
7.60
7.40
7.20
7.00
δ (ppm)

confirmed to be the Hβ's and Hγ of a proline ring (Pro71, see Chapter 6 for details). No other strong NOEs were detected, thus demonstrating that the above NOEs are not due to spin diffusion. Upon selective saturation the heme 5-CH_3 resonance, which partially overlaps the resonance at 19.3 ppm, no NOEs were observed on any of the identified Tyr67 or Pro71 signals, thus ensuring that they originate from the exchangeable signal. On the basis of these data, this spin pattern was assigned to the Tyr67 ring and the signal at 19.3 ppm to its exchangeable hydroxyl proton.

S. c. CN-Fe(III)WTcyt c

A proton signal with properties similar to those of the CN-Ala80 Tyr67 Hη (hydroxyl proton) resonance was identified at 22.2 ppm in the 200 MHz ^{1}H NMR spectrum of CN-Fe(III)WT (Figure 5.11A) taken at 288 K. A superWEFT spectrum taken with short τ and recycle times of 10 ms enhances the intensity of this signal, as seen in Figure 5.11B. By analogy to the results obtained for the mutant, this signal is assigned to the Tyr67 hydroxyl proton.

Assignment of His18

The spectra obtained for CN-Fe(III)Ala80cyt *c* by the superWEFT pulse sequence with different repetition times and τ delays are shown in Figure 5.12. This procedure[9-11] allowed the detection of two broad signals at 16.1 and -3.4 ppm.[24] As reported for the cyanide adducts of peroxidases and Mb, these signals can be attributed to the nonexchangeable ring protons (Hδ2 and Hε1) of the axial histidine ligand.[2] Saturation of the upfield signal at 200 MHz (because it is sharper than at 600 MHz) allows the detection of an NOE with an exchangeable proton that resonates at 16.6 ppm. The presence of this dipolar connectivity is consistent with the assignment of the broad upfield signal as Hε1 and the exchangeable signal as Hδ1 of His18. Analysis of the NOESY and COSY spectra (Figure 5.13) led to the assignment of all of the

Figure 5.11 200 MHz ^{1}H NMR spectra of *S. c.* CN-Fe(III)WTcyt *c* (288 K, 50 mM NaP_i (uncorrected pH 7.0), H_2O) recorded by using the superWEFT pulse sequence with different recycle times and τ values. (A) 250 ms recycle time, 250 ms τ, (B) 10 ms recycle time, 10 ms τ. The broad resonances at 27 and -13 ppm are due to the Hδ2 and Hε1 of the proximal His. The broad resonance at 22.2 ppm is assigned to the Tyr67 OH proton.

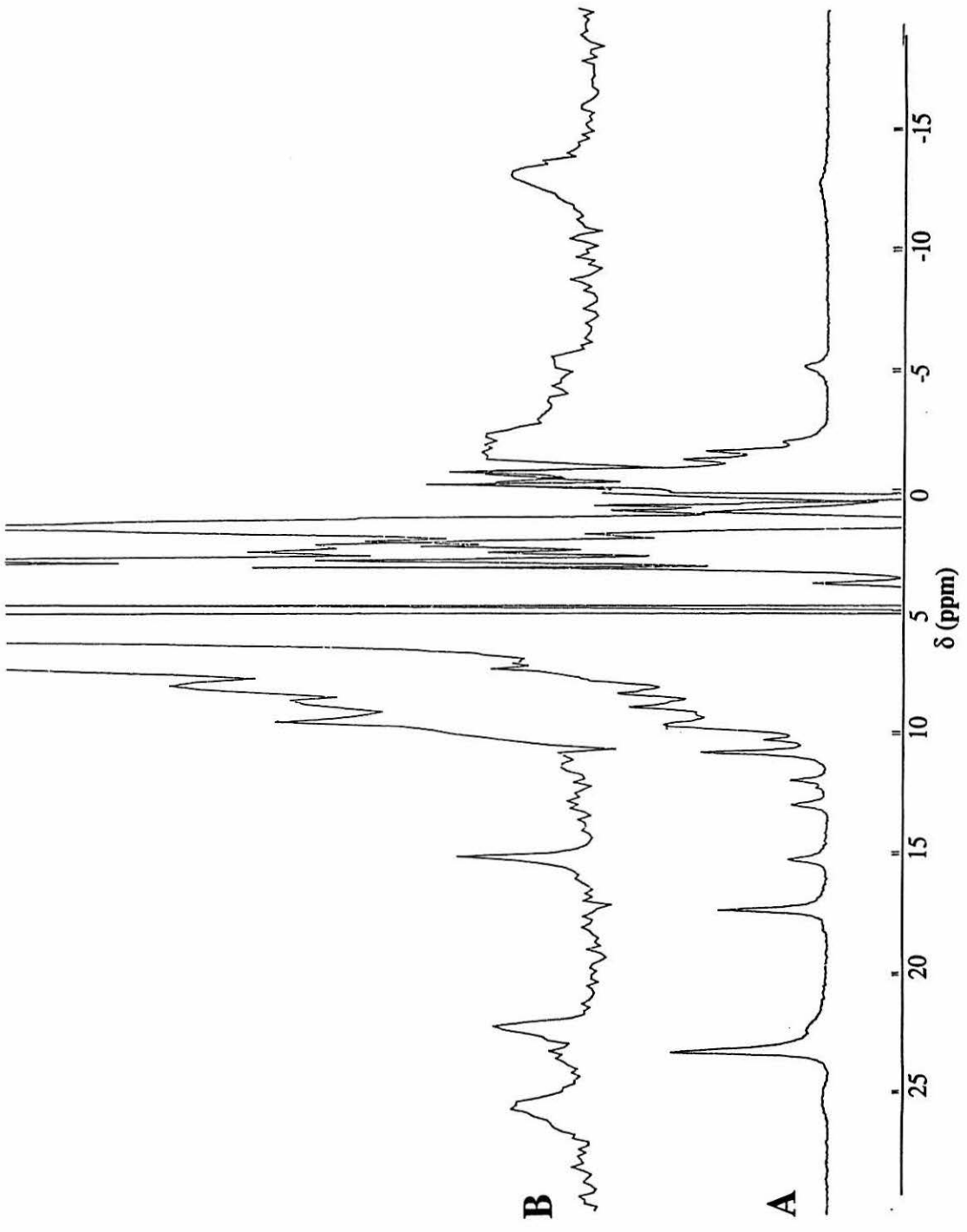

B
A
25
20
15
10
5
0
-5
-10
-15
δ (ppm)

Figure 5.12 200 MHz ^{1}H NMR spectra of *S. c.* CN-Fe(III)Ala80cyt *c* (303 K, 50 mM NaP_i (uncorrected pH 7.0), D_2O) recorded by using the superWEFT pulse sequence with different recycle times and τ values. (A) 250 ms recycle time, 250 ms τ, (B) 33 ms recycle time, 20 ms τ, (C) 10 ms recycle time, 10 ms τ. The broad resonances at 16.1 and -3.4 ppm are due to the Hδ2 and Hε1 of the proximal His.

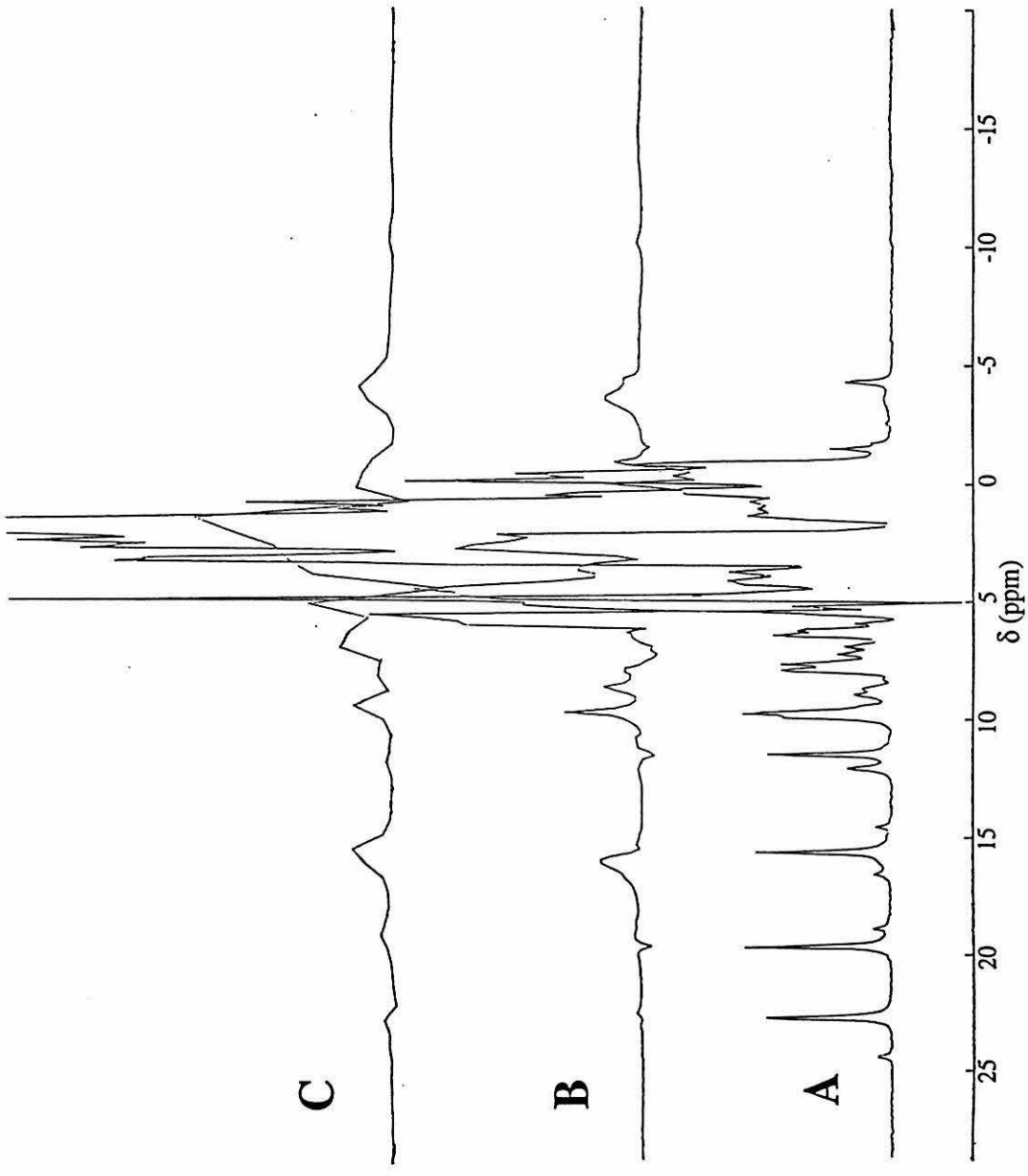
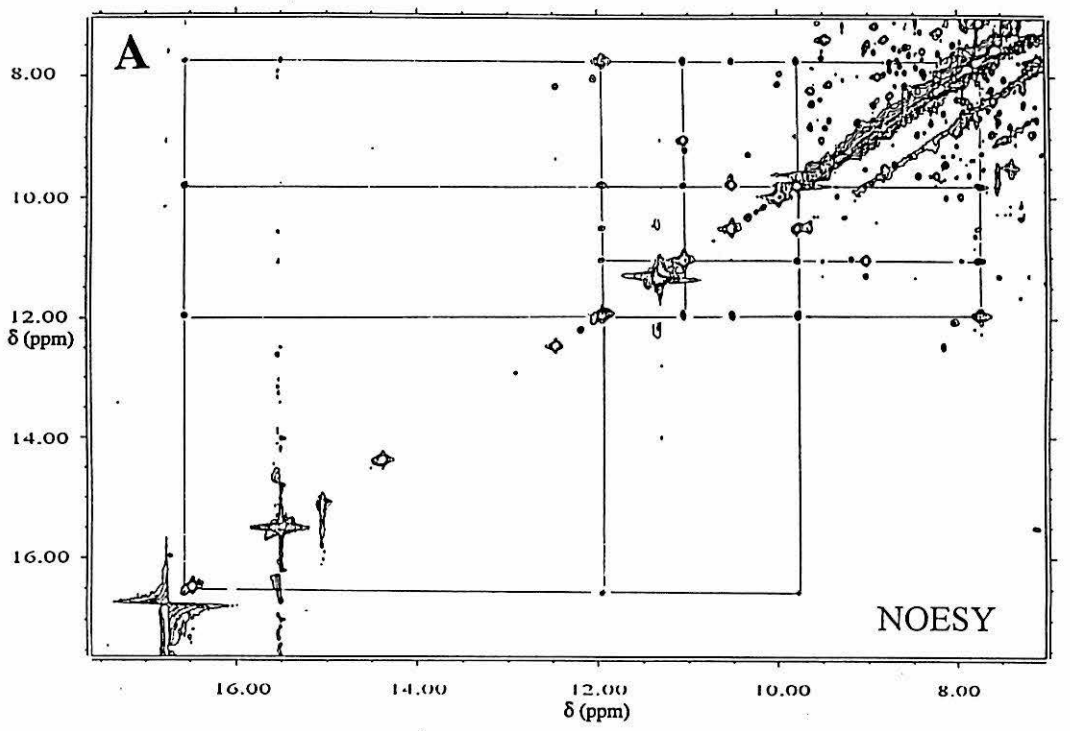

C
B
A
25
20
15
10
5
δ (ppm)
0
-5
-10
-15

Figure 5.13 600 MHz ^{1}H NMR spectra of *S. c.* CN-Fe(III)Ala80cyt *c* (303 K, 50 mM NaP_i (uncorrected pH 7.0)) showing connectivities between the His18 resonances: (A) TPPI NOESY recorded in H_2O with presaturation of the water signal during the relaxation delay (800 ms) and mixing time (100 ms), (B) Magnitude COSY recorded in D_2O with presaturation of the residual water signal during the relaxation delay (800 ms).


A
8.00
10.00
12.00
δ (ppm)
14.00
16.00
NOESY
16.00
14.00
12.00
10.00
8.00
δ (ppm)

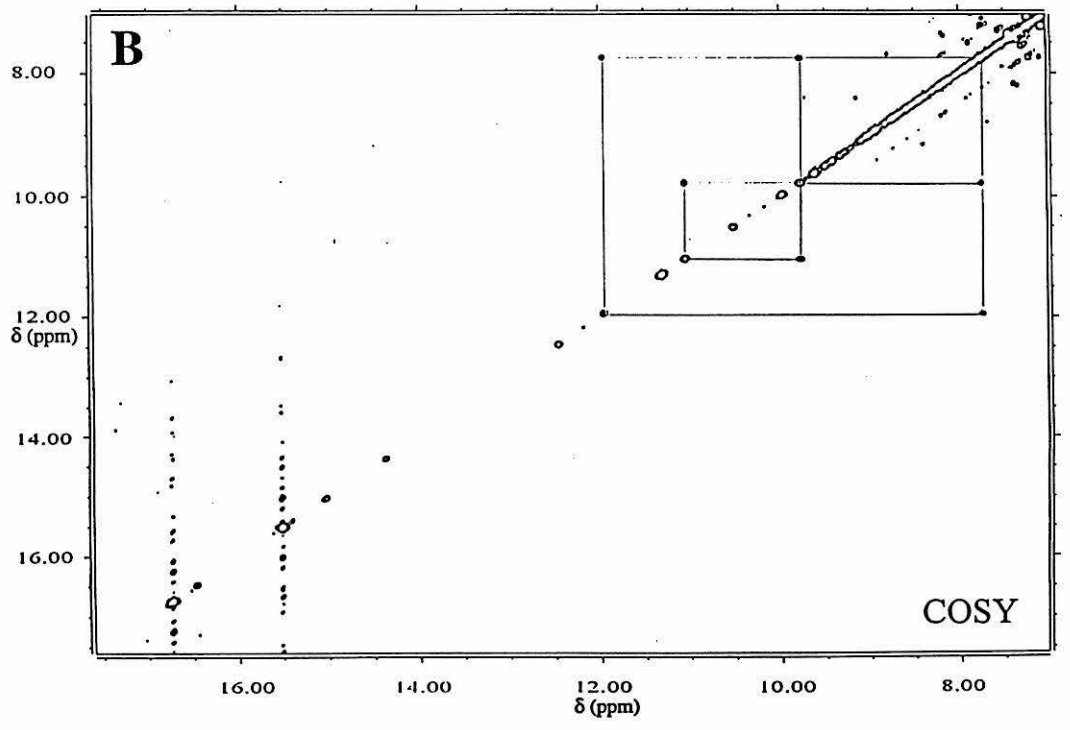

B
8.00
10.00
12.00
δ (ppm)
14.00
16.00
COSY
16.00
14.00
12.00
10.00
8.00
δ (ppm)

Table 5.2

Assignments for the ^1H NMR resonances of His18 and Tyr67 for the cyanide adducts of Fe(III)Ala80cyt c and Fe(III)WTcyt c[a]

Assignment	CN-Ala80cyt *c*, ppm	CN-WTcyt *c*, ppm
His 18		
NHp	11.0	10.6
Hα	9.8	9.6
Hβ	12.0	12.0
Hβ'	7.7	8.8
Hδ1	16.6	15.0
Hδ2	16.1	22.4
Hε1	−3.4	−11.4
Tyr67		
NHp	8.8	*
Hα	4.5	*
Hβ	3.3	*
Hβ'	3.5	*
Hδ1	7.2	*
Hδ2	7.3	*
Hε1	7.8	*
Hε2	7.6	*
Hη	19.3	22.2[b]

[a]T = 303 K unless noted otherwise.

[b]T = 288 K.

*assignment not attempted.

remaining protons of His18 (see Table 5.2). Specifically, NOESY cross peaks have been detected between His18 Hδ1 and the three protons of the AMX spin-system of His18, and among the protons of the AMX system. Scalar connectivities defining the His18 AMX (α/β protons) spin pattern were observed in TOCSY and COSY (Figure 5.13B) maps. The signals of the proximal His in CN-Fe(III)WTcyt *c* have been assigned by the same

procedure. The superWEFT spectrum that aided in the identification of the His18 ring protons is shown in Figure 5.11B. The shift values are reported in Table 5.2.

EPR Spectroscopy

The **g** values measured for *S. c.* CN-Fe(III)Ala80cyt *c* at 6 K are $g_x = 1.2$, $g_y = 1.99$, $g_z = 3.30$. The **g** values for *S. c.* CN-WTcyt *c* at 6 K are $g_x = 1.2$, $g_y = 2.03$, and $g_z = 3.23$.

Discussion

Analysis of the Heme Electronic Structure

The contact shift contributes significantly to the hyperfine shift for heme methyls in low-spin iron(III) systems.[2,3,25,26] Since the unpaired electron of low-spin Fe(III) resides in a d_{xz} or d_{yz} orbital, π unpaired electron spin density is expected on all heme substituents except for the meso protons, while σ spin delocalization should not occur. The heme $e(\pi)$ orbitals (the HOMOs) that mix with the iron d_{xz} and d_{yz} orbitals have nodes at the meso positions. As a result, the meso protons experience hyperfine shifts that originate predominantly from the pseudocontact term.[25,26] The orientation of the heme axial ligands strongly influences both the contact and pseudocontact terms, and is the most important factor in determining the heme methyl and meso proton shifts. The interactions of His and Met donors in the axial positions of the heme remove the degeneracy of the d_{xz} and d_{yz} orbitals, so that the unpaired electron is only in the upper energy orbital; this gives rise to a characteristic spin-density pattern on the heme.[2,26-28] When the z axis of the magnetic susceptibility tensor is essentially perpendicular to the heme plane, as is often found for heme proteins,[14a,29] the four meso protons all experience the same upfield shift given by the axial term of the pseudocontact contribution (see equation 4.3). The angular dependence of the rhombic contribution to

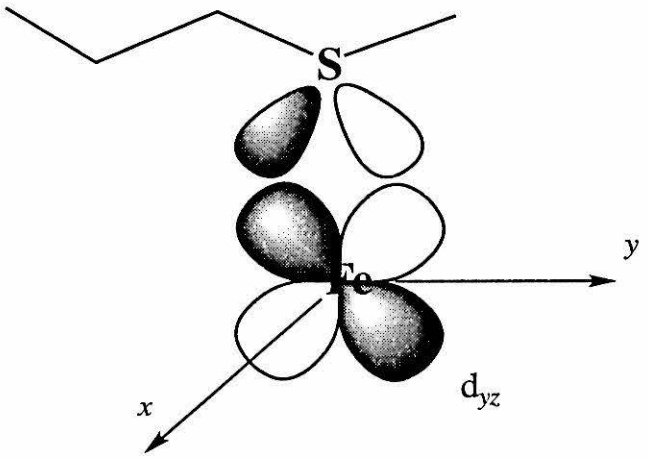
the pseudocontact term causes alternation between adjacent meso proton shifts.[26] The experimental results (Table 5.1, 5.2) can be rationalized with this picture.

The orientation of the heme axial ligands is important in the definition of the rhombic axes *x* and *y*.[25-30] In Met-ligated hemes, the sp^3 lone electron pair on the axial Met sulfur interacts with the d_{xy} and d_{yz} orbitals of the iron (Figure 5.14). The orientation of the axial Met therefore is believed to determine the *x* and *y* magnetic axes for *c*-type cytochromes.[27,31] Recent data indicate that the His orientation also has considerable influence on the asymmetry of the cyt *c* heme electronic structure.[32] The NMR spectrum of WTcyt *c* at pH 7 shows a pairwise pattern for the heme methyl shifts (Figure 4.9A, Table 5.2). The two most-shifted methyl resonances (data collected at 303 K) are 8-CH_3 (34.6 ppm) and 3-CH_3 (31.8 ppm), i.e., the methyls on opposite pyrroles (Figure 5.7). The shift values for the other two heme methyl signals are 10.1 (5-CH_3) and 7.1 (1-CH_3). The pairwise effect is also observed on the other pyrrole substituents (Table 5.2). However, analysis of the shift dependence of the protons in position α with respect to the heme, except for the heme methyls which freely rotate, is complicated by the dihedral-angle dependence.[33] This analysis therefore concentrates on the methyl resonances.

Upon cyanide binding to WTcyt *c*, Met80 is detached from the iron and the heme methyl shift pattern changes completely, being determined only by the orientation of the proximal His ligand. An almost equivalent contact-shift contribution for the four heme methyls is observed, probably due to the orientation of the His ring plane along the α- and γ-meso positions (Figure 5.15).[30] In this case, the heme methyl resonances have a mean shift value of 17.9 and a difference of only 11.1 ppm between the one most upfield and the one most downfield, compared to a difference of 27.5 ppm seen for WTcyt *c* without bound cyanide. In the case of CN-Fe(III)Mb, the His plane is oriented along pyrroles I and III (Figure 5.15) and provides large shifts for methyls 1 and 5. In the case

Figure 5.14 Representation of interactions between iron d orbitals and the orbitals of axial Met and His ligands. (A) The lone pair of the Met is shown interacting with an iron d_{yz} orbital. (B) A His π orbital interacting with a metal d_{xz} orbital. These interactions remove the degeneracy of the iron d_{yz} and d_{xz} orbitals, leading to the unpaired electron spin density pattern on the heme observed by paramagnetic NMR. Note that the x and y axes are interchanged between (A) and (B).

A



B

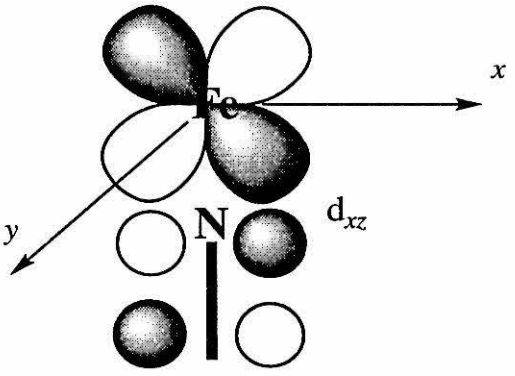
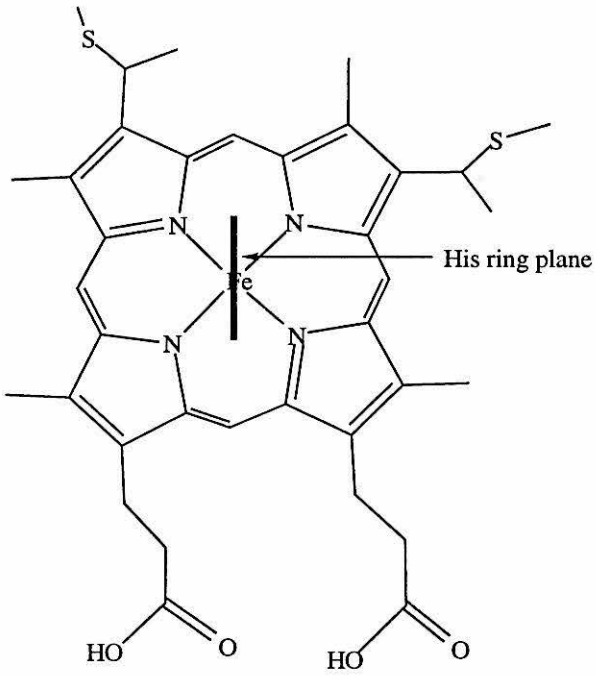


Figure 5.15 Orientation of His ligand in (A) cytochrome *c* and (B) Mb. When the second axial ligand is CN, the orientation of the iron-bound His determines the asymmetry of the unpaired electron spin density on the heme and, thus, the pattern of hyperfine shifts observed in the NMR spectrum of the protein.

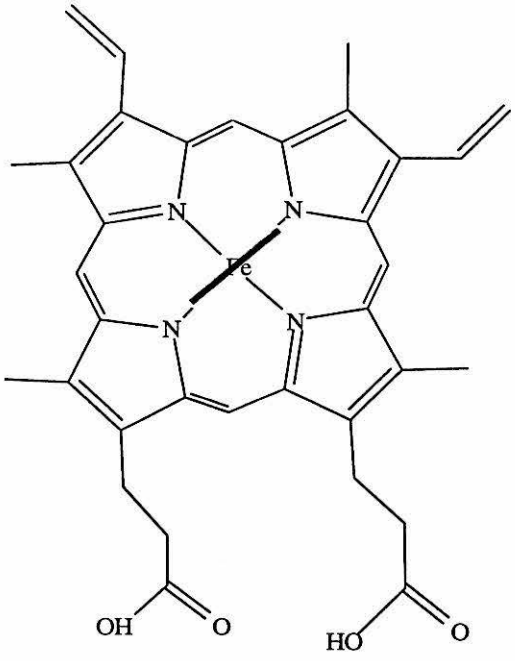
A

cyt *c*



B

Mb



of peroxidases, the proximal His is oriented along pyrroles II and IV (C*c*P)[34] or along the α- and γ-meso positions (lignin peroxidase).[35] In both cases, large shifts are observed for heme methyls 3 and 8.[2]

For the ^{1}H NMR resonances of the CN-Fe(III)Ala80cyt *c* heme methyls, a mean shift of 17.1 and a difference of 11.2 ppm, similar to the values of CN-Fe(III)WTcyt *c*, are observed. This indicates that the proximal His orientation[22] is maintained in the variant. A minor difference between the heme methyl shifts for CN-Ala80 and CN-WT is seen for 1-CH_3 and 5-CH_3, which may be accounted for by assuming small differences in the pseudocontact contribution. Consistent with this hypothesis, a slightly different shift for the δ-meso proton of CN-Ala80cyt *c* compared to CN-WTcyt *c* is observed. In addition, the His18 Hε1 and Hδ2 ring protons have significantly different shifts; these resonances have been shown to be very sensitive to differences in the orientation of the *z* axis of the magnetic susceptibility tensor.[14a,29] To detect this difference in the systems studied here, the **g** values for CN-WTcyt *c* and CN-Ala80cyt *c* have been compared but were found to be nearly identical (see Results); thus, any difference in the pseudocontact contribution between the two proteins must be due to differences in the orientation of the magnetic susceptibility tensor and not in its magnitude.

The temperature dependence of the hyperfine shifts is related to the energy separation between the ground and excited electronic states which, in turn, is modulated by interactions between the iron and its ligands. For low-spin WTcyt *c*, it has been found that the shifts of methyls 1 and 5 increase as temperature increases (anti-Curie behavior), while the shifts of methyls 3 and 8 decrease as temperature increases (Curie behavior).[36] This anti-Curie effect has been interpreted in terms of a Boltzmann distribution between two partially filled d_{xz}- and d_{yz}-symmetry molecular orbitals, assuming a splitting of 3 kJ mol^{-1}.[37] A similar anti-Curie effect was not observed for the Ala80 and WT cyanide adducts.

Characteristics of the Tyr67 Hydroxyl Proton Resonance

Tyr67 has been proposed to play a role in the stabilization of the dioxygen adducts of h. h. and *S. c.* Fe(II)Ala80cyt *c* by hydrogen-bonding with the heme-O_2 unit. To explore the validity of this hypothesis, the assignment of the Tyr67 protons is needed; in particular, the analysis of the Tyr ring hydroxyl shift, relaxation properties, and exchange rate would provide an indication of its distance from the heme and the possibility of it being involved in a hydrogen-bonding interaction with the heme-bound cyanide. Through a combination of 1D NOE and 2D TOCSY (Figure 5.10) and NOESY experiments employing short mixing times, the ring protons, including the hydroxyl proton (Hη), of Tyr67 have been assigned. The Hη proton signal is visible in the 1D ^{1}H NMR spectrum of CN-Fe(III)Ala80 in H_2O as a broad resonance at 19.3 ppm (inset of Figure 5.2). One striking feature of the Tyr67 Hη signal is its large hyperfine shift value (compared to the typical Tyr Hη shift of 10 ppm).[38] This shift indicates that this proton experiences significant hyperfine coupling with the iron. In addition, this nucleus has an extremely short T_1 value of 9 ± 2 ms. Comparison of the T_1 value with that of the Hε1 proton of His18, which is 3.3 Å from the metal ion in WTcyt *c*, suggests that this Tyr67 hydroxyl proton is 4 ± 0.2 Å from the iron.[39]

The fact that the Tyr67 resonance can be observed at all is significant. Tyr Hη protons usually exchange too fast with bulk water to be observed.[38] However, if the Tyr Hη is involved in a hydrogen-bond, this may significantly slow its exchange rate, allowing its detection. The corresponding Tyr67 hydroxyl proton signal in CN-WTcyt *c* is observed only at lower temperatures (288 K or below), indicating that this proton is exchanging faster than in the mutant.

In cyanide derivatives of metMb and peroxidases, the broad resonance of a proton with similar relaxation properties and slow exchange with bulk water has been assigned to a labile distal His ring proton that is believed to be hydrogen-bonded with the cyanide

nitrogen.[12,14ej,40] The data here suggest that, in analogy with the metMb and peroxidase systems, the Tyr67 hydroxyl hydrogen-bonds to the cyanide nitrogen (Figure 5.16). This is confirmed by NOEs from the Tyr protons that allow the location of the residue in the 3D structure (discussed in Chapter 6). The presence of hydrogen-bonding between the Tyr67 hydroxyl proton and the cyanide nitrogen suggests that a similar interaction may occur between Tyr67 and a heme-bound O_2 in the reduced mutant, analogous to the hydrogen-bonding involving the distal histidine in the dioxygen adduct of Mb.

References and Notes

(1) The work in this chapter was done at the University of Florence in collaboration with Dr. Paola Turano, Prof. Lucia Banci, and Prof. Ivano Bertini.

(2) Bertini, I.; Turano, P.; Vila, A. J. *Chem. Rev.* **1993,** *93,* 2833-2932.

(3) Bertini, I.; Luchinat, C. *NMR of Paramagnetic Molecules in Biological Systems*; Benjamin/Cummings: Menlo Park, 1986.

(4) McConnell, H. M.; Robertson, R. E. *J. Chem. Phys.* **1958,** *29,* 1361-1365. Kurland, R. J.; McGarvery, B. R. *J. Magn. Res.* **1970,** *2,* 286-304.

(5) McConnell, H. M.; Chestnut, D. B. *J. Chem. Phys.* **1958,** *28,* 107-117. McConnell, H. M. *J. Chem. Phys.* **1956,** *24,* 764-766.

(6) Horrocks, W. D., Jr.; Hall, D. D. *Inorg. Chem.* **1971,** *10,* 2368- 2370. Vega, A. J.; Fiat, D. *Mol. Phys.* **1976,** *31,* 347-355.

(7) Koenig, S. H. *J. Magn. Reson.* **1982,** *47,* 441-453. Bloembergen, N. *J. Chem. Phys.* **1957,** *27,* 572-573.

(8) Satterlee, J. D.; Erman, J. E.; Mauro, J. M.; Kraut, J. *Biochemistry* **1990,** *29,* 8797-8804.

(9) Ferrer, J. C.; Turano, P.; Banci, L.; Bertini, I.; Morris, E. K.; Smith, K. M.; Smith, M.; Mauk, A. G. *Biochemistry* **1994,** *33,* 7819-7829.

(10) Banci, L.; Bertini, I.; Pease, E. A.; Tien, M.; Turano, P. *Biochemistry,* **1992,** *31,* 10009-10017.

(11) Banci, L.; Bertini, I.; Turano, P.; Tien, M.; Kirk, T. K. *Proc. Natl. Acad. Sci. U.S.A.* **1991,** *88,* 6956-6960.

(12) Banci, L.; Bertini, I.; Turano, P.; Ferrer, J. C.; Mauk, A. G. *Inorg. Chem.* **1991,** *30,* 4510-4516.

(13) Banci, L.; Bertini, I.; Kuan, I-C.; Tien, M.; Turano, P.; Vila, A. J. *Biochemistry* **1993,** *32,* 13483-13489.

(14) (a) Rajarathnam, K.; La Mar, G. N.; Chiu, M. L.; Sligar, S. G. *J. Am. Chem. Soc.* **1992,** *114,* 9048-9058. (b) de Ropp, J. S.; La Mar, G. N.; Wariishi, H.; Gold, M. H. *J. Biol. Chem.* **1991,** *266,* 15001-15008. (c) de Ropp, J. S.; Yu, L. P.; La Mar, G. N. *J. Biomol. NMR* **1991,** *1,* 175-190. (d) de Ropp, J. S.; La Mar, G. N. *J. Am. Chem. Soc.* **1991,** *113,* 4348-4340. (e) Thanabal, V.; La Mar, G. N. *Biochemistry* **1989,** *28,* 7038-7044. (f) Emerson, S. D.; Lecomte, J. T. J.; La Mar, G. N. *J. Am. Chem. Soc.* **1988,** *110,* 4176-4182. (g) Thanabal, V.; de Ropp, J. S.; La Mar, G. N. *J. Am. Chem. Soc.* **1988,** *110,* 3027-3035. (h) Thanabal, V.; de Ropp, J. S.; La Mar, G. N. *J. Am. Chem. Soc.* **1987,** *109,* 265-272. (i) Thanabal, V.; de Ropp, J. S.; La Mar, G. N. *J. Am. Chem. Soc.* **1987,** *109,* 7516-7525. (j) Lecomte, J. T. J.; La Mar, G. N. *Eur. Biophys. J.* **1986,** *13,* 373-381. (k) de Ropp, J. S.; La Mar, G. N.; Smith, K. M.; Langry, K. C. *J. Am. Chem. Soc.* **1984,** *106,* 4438-4444. (l) La Mar, G. N.; Davis, N. L.; Parish, D. W.; Smith, K. M. *J. Mol. Biol.* **1983,** *168,* 887-896. (m) La Mar, G. N.; de Ropp, J. S.; Chacko, V. P.; Satterlee, J. D. *Biochim. Biophys. Acta* **1982,** *708,* 317-325.

(15) Satterlee, J. D.; Erman, J. E. *Biochemistry* **1991,** *30,* 4398-4405. Satterlee, J. D.; Erman, J. E.; de Ropp, J. S. *J. Biol. Chem.* **1987,** *262,* 11578-11583. Satterlee, J. D.; Erman, J. E.; La Mar, G. N.; Smith, K. M.; Langry, K. C. *J. Am. Chem. Soc.* **1983,** *105,* 2099-2104. Satterlee, J. D.; Erman, J. E. *J. Biol. Chem.* **1983,** *258,* 1050-1056.

(16) Inubushi, T.; Becker, E. D. *J. Magn. Reson.* **1983,** *51,* 128-133.

(17) Banci, L.; Bertini, I.; Luchinat, C.; Piccioli, M.; Scozzafava, A.; Turano, P. *Inorg. Chem.* **1989,** *28,* 4650-4656. Unger, S. W.; Lecomte, J. T. J.; La Mar, G. N. *J. Magn. Reson.* **1985,** *64,* 521-526.

(18) Macura, S.; Wüthrich, K.; Ernst, R. R. *J. Magn. Reson.* **1982,** *47,* 351-357. Marion, D.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983,** *113,* 967-974.

(19) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985,** *65,* 355-360. Griesinger, C.; Otting, G.; Wüthrich, K. Ernst, R. R. *J. Am. Chem. Soc.* **1988,** *110,* 7870-7872.

(20) Bax, A.; Freeman, R.; Morris, G. *J. Magn. Reson.* **1981,** *42,* 164-168. Bax, A.; Freeman, R. *J. Magn. Reson.* **1981,** *44,* 542-561.

(21) Eccles, C.; Güntert, P.; Billeter, M.; Wüthrich, K. *J. Biomol. NMR* **1991,** *1,* 111-130.

(22) Louie, G. V.; Hutcheon, W. L. B.; Brayer, G. D. *J. Mol. Biol.* **1988,** *199,* 295-314. Louie, G. V.; Brayer, G. D. *J. Mol. Biol.* **1990,** *214,* 527-555.

(23) Gao, Y.; Boyd, J.; Williams, R. J. P.; Pielak, G. J. *Biochemistry* **1990,** *29,* 9664-7003.

(24) The odd shapes of the signals in Figure are due to the low digital resolution required by the fast recycle time of the experiment. Zero-filling would provide better-shaped signals.

(25) La Mar, G. N.; Walker, F. A. In *The Porphyrins*; Dolphin, D., Ed.; Academic Press: New York, 1979; pp 61-157.

(26) Lee, K. -B.; La Mar, G. N.; Mansfield, K. E.; Smith, K. M.; Pochapasky, T. C.; Sligar, S. G. *Biochim. Biophys. Acta* **1993,** *1202,* 189-199.

(27) Satterlee, J. D. *Annu. Rep. NMR Spectrosc.* **1986,** *17,* 79-178.

(28) Traylor, T. G.; Berzinis, A. P. *J. Am. Chem. Soc.* **1980,** *102,* 2844-2846. Keller, R. M.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1978,** *83,* 1132-1139. La Mar, G. N.; Viscio, D. B.; Smith, K. M.; Caughey, W. S.; Smith, M. L. *J. Am. Chem. Soc.* **1978,** *100,* 8085-8092. Shulman, R. G.; Glarum, S. H.; Karplus, M. *J. Mol. Biol.* **1971,** *57,* 93-115.

(29) Emerson, S. D.; La Mar, G. N. *Biochemistry* **1990,** *29,* 1556-1566.

(30) Walker, F. A.; Benson, M. *J. Phys. Chem.* **1982,** *86,* 3495-3499. Walker, F. A. *J. Am. Chem. Soc.* **1980,** *102,* 3254-3256.

(31) Senn, H.; Böhme, H.; Wüthrich, K. *Biochim. Biophys. Acta* **1984,** *789,* 311-323, and references therein.

(32) Turner, D. L. *Eur. J. Biochem.* **1995,** *227,* 829-837.

(33) Bertini, I.; Turano, P. In *NMR of Paramagnetic Macromolecules. NATO ASI Series*; La Mar, G. N., Ed.; Kluwer Academic: Dordrecht, 1994, and references therein.

(34) Finzel, B. C.; Poulos, T. L.; Kraut, J. *J. Biol. Chem.* **1984,** *259,* 3027-3036.

(35) Edwards, S. L.; Raag, R.; Wariishi, H.; Gold, M. H.; Poulos, T. L. *Proc. Natl. Acad. Sci. U.S.A.* **1993,** *90,* 750-754. Poulos, T. L.; Edwards, S. L.; Wariishi, H.; Gold, M. H. *J. Biol. Chem.* **1993,** *268,* 4429-4440. Pintek, K.; Glumoff, T.; Winterhalter, K. *FEBS Lett.* **1993,** *315,* 119-124.

(36) Senn, H.; Wüthrich, K. *Biochim. Biophys. Acta* **1983,** *743,* 69-81.

(37) Turner, D. L. *Eur. J. Biochem.* **1993,** *211,* 563-568.

(38) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley-Interscience: New York, 1986.

(39) When nonselective T_1 values are lower than 100 ms, the recovery of the magnetization after a nonselective excitation shows exponential behavior which, therefore, can be fitted to a single exponential (see Figure 4.8). This means that

the dominant contribution to nuclear relaxation is the coupling between the resonating nucleus and the unpaired electron, and it is therefore related to the inverse sixth power of the metal-nucleus distance. Since the T_1 of the bulk water is ~ 0.70 s, the system is in the slow exchange limit. Also, because the T_1 is nonselective, the effect that any exchange may have on the T_1 of the Hη signal does not alter the determination of the iron-proton distance beyond the uncertainty given above.

(40) Cutnell, J. D.; La Mar, G. N.; Kong, S. B. *J. Am. Chem. Soc.* **1981,** *103,* 3567-3572.

Chapter 6

The Three-Dimensional Solution Structure of the Cyanide Adduct of *Saccharomyces cerevisiae* Ala80-iso-1-cytochrome c[1]

Background

For nearly 25 years, X-ray crystallography stood as the sole method for determining three-dimensional structures of proteins at atomic resolution.[2] That changed in 1984, when the structure of proteinase inhibitor-IIA from bull seminal plasma was solved by ^{1}H NMR spectroscopy and distance geometry (DG) calculations, demonstrating that NMR is a tractable method for the *ab initio* determination of protein structures in solution.[3] As of the end of 1994, over 250 solution structures of biological macromolecules had been solved by NMR,[4] and the number of new structures reported continues to increase.

NMR spectroscopy is an attractive method for protein structure determination because it allows the investigation of the protein in its natural environment, *i.e.* in solution. In most cases where structures of the same protein have been solved both in the solution and the crystalline states, it has been comforting to learn that the structures agree remarkably well.[5] However, there are a few cases of minor, and even major discrepancies between crystal and solution structures; in some cases these have been attributed to errors in structure determination, and in a few they are due to real differences in molecular structure.[5] NMR therefore has served as an important check of X-ray crystal structures and whether it is reasonable to apply structural information obtained on a crystalline protein sample to the solution state.

Besides the ability to study molecular structure in solution, NMR has the advantage of allowing the investigation of internal mobility and labile proton exchange rates. Techniques have also been developed to study the exchange of interior hydration water molecules with bulk solvent.[6] Of course, it also offers a means to determine the structure of proteins for which suitable crystals cannot be obtained. The main drawback of structure determination by NMR is that it is feasible only for relatively small proteins (> 200 amino acids). For proteins > 100 amino acids, it often is necessary to isotopically label the protein and obtain three- and four- dimensional spectra to alleviate poor spectral

resolution due to broad signals and greater spectral overlap.[7] This of course requires the overexpression of the protein and can require both significant time and expense.

An additional challenge in the field of structure determination by NMR is presented by the presence of paramagnetic centers in some metalloproteins.[8,9] The resulting reduction in spectral quality due to the broadening of resonances due to nuclei near the metal is particularly disconcerting because the paramagnetic center is usually at the protein active site and, thus, the most important area of a metalloprotein for the study of structure/function relationships. However, developments in NMR methodology[9,10] have allowed extensive sequence-specific assignments on paramagnetic metalloproteins and, in 1994, the first solution structure of a paramagnetic metalloprotein (*Ectothiorhodospira halophila* high-potential iron-sulfur protein I) was reported.[11] Since then, the structures of three other paramagnetic iron-sulfur proteins have been published, ranging in size from 55 to 85 amino acids.[12]

In this chapter, the sequence-specific ^{1}H NMR assignment and resulting solution structure of the cyanide adduct of *S. c.* Fe(III)Ala80cyt *c* are presented. At 108 amino acids, this is the largest paramagnetic protein for which an NMR structure has been determined. The ways in which the paramagnetism both helped and hindered the sequence-specific assignment are discussed. The relationship between the solution structure of *S. c.* CN-Fe(III)Ala80cyt *c* and the stability of O_2-Fe(II)Ala80cyt *c* is discussed.

Materials and Methods

^{1}H NMR Spectroscopy

Expression and purification of *S. c.* Gln39Ala80Ser102iso-1-cytochrome *c* (*S. c.* Ala80cyt *c*) was described in Chapter 3. ^{1}H NMR sample preparation was performed as indicated in Chapter 4 except that all samples were 3 mM in protein. The methodology of collecting 1D ^{1}H NMR experiments at 200 MHz and 600 MHz with solvent

suppression was given in Chapter 5. The collection and processing of 1D NOE and 2D NOESY, COSY, TOCSY, and NOE-NOESY experiments were also described in Chapter 5. The diamagnetic regions of the TOCSY and NOESY spectra of *S. c.* CN-Fe(III)Ala80cyt *c* acquired in D_2O are pictured in Figure 6.1.

Distance Geometry Calculations

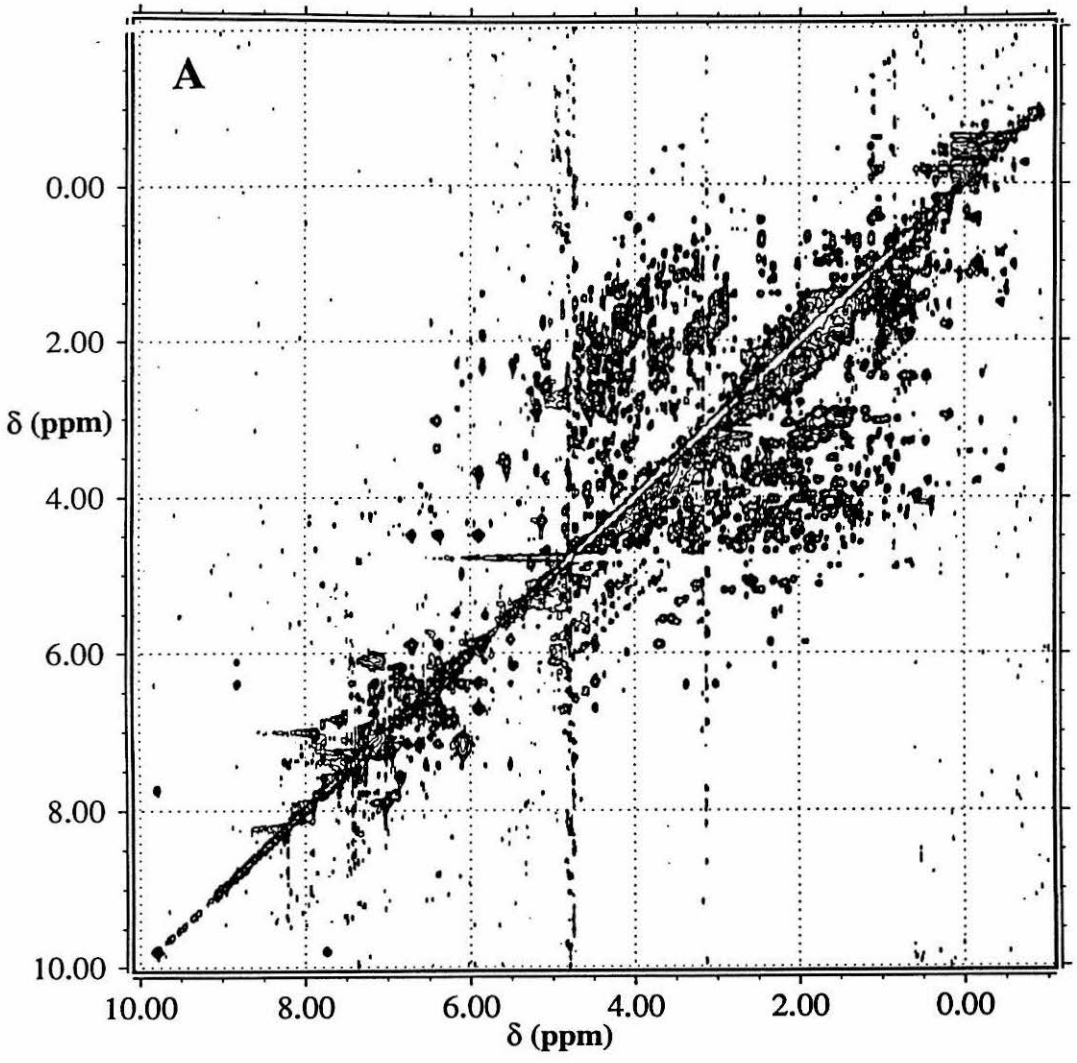
The volumes of the cross peaks between assigned resonances were obtained using the integration routines in the program XEASY. Most of the dipolar connectivities were taken from the 100-ms mixing-time NOESY in H_2O solution at 303 K. Connectivities whose volumes could be better measured in other spectra recorded with the same parameters, but either at a different temperature or in D_2O solution, were scaled referring to a few intraresidue connectivities whose volumes could be accurately measured in both spectra.

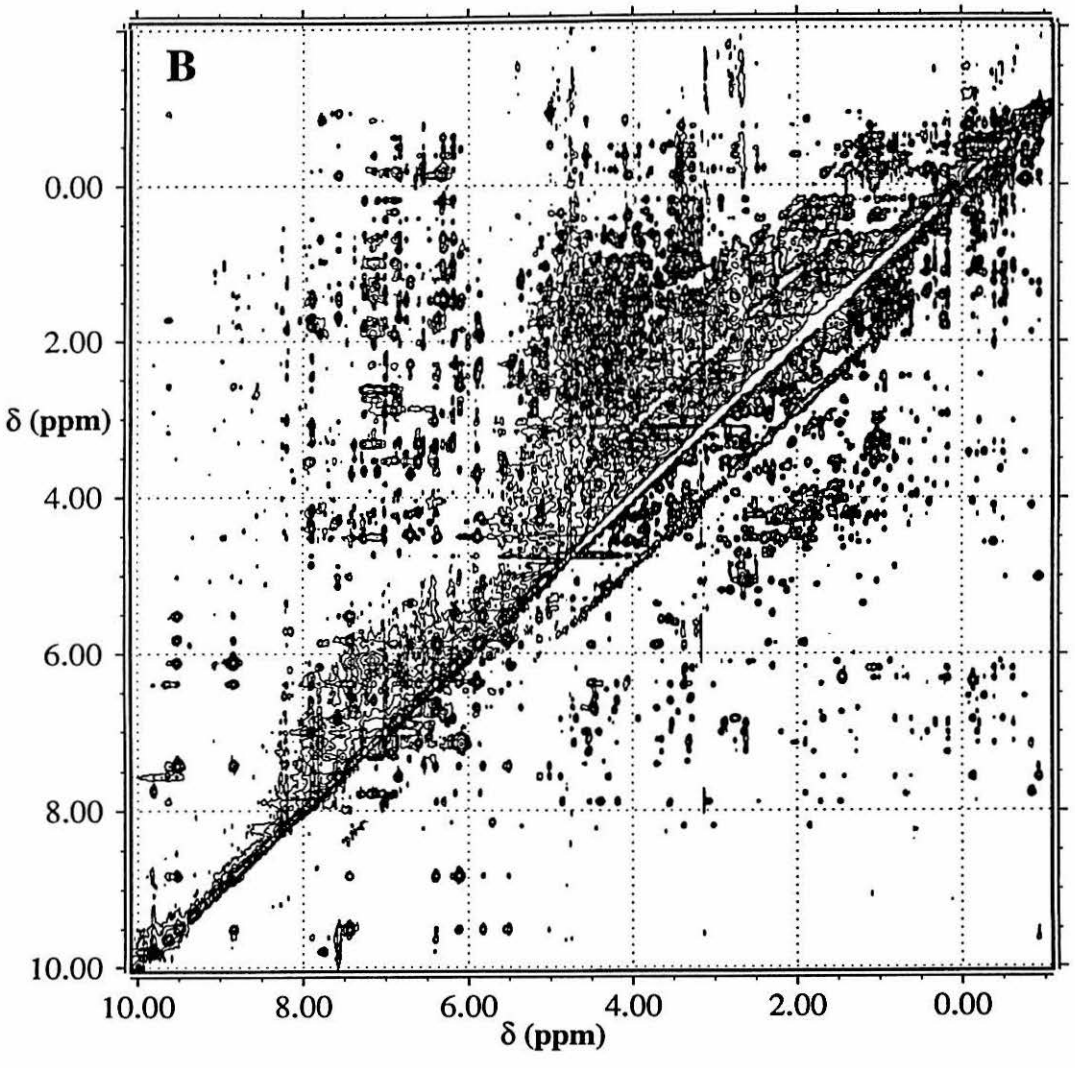
NOESY cross peak intensities were converted into upper limits of interatomic distances by the use of the program CALIBA.[13] As previously described by Wüthrich *et al.*,[13-16] the upper limits obtained form CALIBA were classified as *i*) intraresidue except NH, Hα, Hβ; *ii*) sequential and intraresidue NH, Hα, Hβ; *iii*) medium-range (all non-sequential interresidue connectivities between NH, Hα, Hβ within a segment of five consecutive residues); *iv*) long-range backbone; *v*) long-range. For NOEs involving methyl groups, the upper limits were evaluated independently using five analogous classes. The experimental distance constraints were then used to generate protein conformers by using the distance geometry (DG) program DIANA.[13]

Upper and lower distance limits were imposed to build up the heme. In addition, the non-standard amino acid trimethyllysine (Tml) present at position 72 was built and added to the DIANA library.

With the preliminary structures available, the scaling factors for the volume-to-distance conversion for each class were evaluated by plotting volumes of peaks arising

Figure 6.1 600 MHz ^{1}H NMR (A) TOCSY and (B) NOESY spectra of *S. c.* CN-Fe(III)Ala80cyt *c* at 303 K. The 3-mM protein sample was in 50 mM NaP_i (uncorrected pH 7.0) in D_2O.


A
0.00
2.00
δ (ppm)
4.00
6.00
8.00
10.00
10.00
8.00
6.00
4.00
2.00
0.00
δ (ppm)


B
0.00
2.00
δ (ppm)
4.00
6.00
8.00
10.00
10.00
8.00
6.00
4.00
2.00
0.00
δ (ppm)

from pairs of protons at fixed distances vs. interatomic distance. As the structure emerged form successive runs of DG calculations, selected interatomic distances were taken from these structures and additional calibrations were performed. Several cycles of the structure calculation were carried out in order to recalibrate the NOE distance constraints; new NOESY cross peaks could now be assigned using the program ASNO.[17]

For the DG calculations, the default minimization parameters were used with the following exceptions: the maximum number of target function evaluations was increased to 3000 at all levels except the final three, where a number of 7000 proved to yield an optimal convergence for the calculated structures. In addition to the standard procedure of the DG calculations, extensive use of the redundant angle strategy (REDAC)[15] was employed for the definition of starting conformations. Angle constraints for dihedral angles were only accepted by the program if in at least 10 structures the local target function for the residue in question was equal to or smaller than 0.5 $Å^2$. During the last four REDAC cycles the target function per residue cutoff criterion was 0.1 $Å^2$ to exclude meaningless angle constraints. The final DG calculation on 39 structures was performed without angle constraints to prove convergence. The program GLOSMA[13] yielded 60 stereospecific assignments for diastereotopic protons (see Results and Discussion).

Restrained Energy Minimization and Restrained Molecular Dynamics

Restrained energy minimization (REM) and restrained molecular dynamics (RMD) calculations on the family of 17 structures with the lowest target functions obtained from DG calculations were performed with the SANDER module of the AMBER 4.0 program package.[18] The force field parameters for all of the residues, except for the heme, the two axial heme ligands, and the heme-bound cysteines, were the standard AMBER "all atom" parameters.[19] The topology file for the heme was adapted from that of a heme *b*, available in the 4.0 release of AMBER and modified in order to take into account the structure of heme *c*. In particular, the two vinyl groups of heme *b*

were modified to $CHCH_3$ moieties. The heme-bound cysteines, Cys14 and Cys17, were obtained from the standard cysteines by removing the sulfhydryl hydrogen and linking the α carbon of the 2-thioether and of the 4-thioether of the heme to each of the cysteines' sulfur atoms with covalent bonds. The force constants for these bonds (stretching, bending and torsion) were taken to be the same as those for the sulfur and ε-CH_3 group of a methionine. The force field parameters, except for the charges, for the heme moiety and for the axial histidine were taken from the standard AMBER database, when available. Alternatively, they were taken to be equal to those already used for molecular dynamics (MD) calculations on peroxidases,[20] which were derived from data obtained on porphyrin model complexes.[21] The parameters for the cyanide and for its bond with the iron were derived from data obtained on small metal-cyanide complexes.[22] The charges of the iron, the heme (including the thioethers), the cyanide, and the axial histidine were determined through semi-empirical calculations using the MOPAC 5.0 program.[23] The AM1 Hamiltonian was used for the self-consistent field calculations. The charges of equivalent atoms in the porphyrin ring were taken as the average of the calculated charges (the difference never being larger than 10% of each point charge). The non-standard amino acid Tml was built with the PREP module of AMBER, and its charge was determined using the MOPAC 5.0 program.

Each structure of the DG family was energy minimized to an RMS energy gradient lower than 0.4 kJ mol^{-1} $Å^{-2}$ with the inclusion of the NOE- and hydrogen-bond-derived distance constraints from REM. RMD calculations were then performed on each structure. The calculations were performed for 48 ps, heating the system from 0 K up to the equilibrium temperature of 300 K. The systems were coupled to a thermal bath at 300 K with a constant of 0.1 ps.[24] The calculations were performed *in vacuo*, with a distance-dependent dielectric constant. The time step used was 1.5 fs; the bond lengths were kept fixed by the SHAKE algorithm[25] during the MD calculations, but not during the energy minimization. The non-bonded interactions were evaluated with a cut-off of 10 Å and the

pair list for the evaluation of the non-bonded interactions was updated every 20 steps. During the REM and RMD calculations the NOE-derived distance constraints were applied through the mixed linear-harmonic potential used by the SANDER module. The harmonic force constant was set to 134 kJ mol^{-1} $Å^{-2}$. The upper limit distances used by DIANA were taken as the upper limit for the application of the harmonic potential.

The final 24 ps of the RMD trajectory for each structure of the family were used to generate average structures with the CARNAL program.[26] These average structures were then subjected to REM calculations.

The DG, REM, and RMD calculations (double precision accuracy) were performed on IBM RISC 6000 computers.

Structure Analysis

Structure analysis, *i.e.* Ramachandran plots, deviation from ideal structural parameters, and secondary structure elements, was performed with the PROCHECK program.[27] Visual inspection and drawings of the various families were performed with the programs SYBYL (Tripos Associates, Inc.), RASMOL,[28] and Insight II (Biosym).

Results and Discussion

Sequence-Specific Assignment

Procedure and Strategy

The sequence-specific assignment began with the identification of amino acid side chain spin systems in TOCSY and COSY maps, as described by Wüthrich.[29] Once as many spin systems as possible were identified, their positions in the sequence were determined through a combination of strategies. Resolved, hyperfine-shifted resonances of heme substituents provided one convenient starting point in the sequence-specific assignment of spin systems. As discussed in Chapter 5, many NOESY cross peaks

Table 6.1

NOESY connectivities observed from heme substituent resonances to amino acids in S. c. CN-Fe(III)Ala80cyt c[a]

heme resonance	amino acid resonance	heme resonance	amino acid resonance
2-Hα	Arg13 Hβ's	7-Hα	Asn52 Hδ1
	Cys14 Hβ1		Trp59 Hε1, Hζ2
2-βCH$_3$	Arg13 Hβ's	7-Hα'	Asn52 Hδ1
	Leu68 δ-CH$_3$'s		Trp59 Hε1, Hζ2, Hη2
	Phe82 Hζ, Hδ/Hε		Tyr67 Hε1
	Leu85 δ-CH$_3$'s	7-Hβ	Leu32 δ-CH$_3$'s
	Leu94 δ-CH$_3$'s		Tyr48 Hη
4-Hα	Cys17 Hβ2		Asn52 Hδ2
	Val28 γ-CH$_3$		Trp59 Hε1, Hζ2
	Ala80 Hα	7-Hβ'	Pro30 Hγ, Hδ
	Ala81 NH		Tyr48 Hε1, Hη
4-βCH$_3$	Gln16 Hγ, Hε's	α-meso	Cys14 Hα
	Val28 γ-CH$_3$		Phe82 Hβ
	Ala81 NH	β-meso	Ala80 Hα
6-Hα	Tyr46 Hε2		Ala80 NH
6-Hα'	Tyr46 Hε2	δ-meso	Leu32 δ-CH$_3$'s
6-Hβ	Lys79 NH		Leu68 δ-CH$_3$'s
	Ala80 NH, Hα, Hβ		Leu98 δ-CH$_3$

[a]connectivities involving heme methyls were indicated in Chapter 5.

between heme methyl groups and heme pocket amino acids were identified; comparing the types of amino acids near each heme methyl to those expected on the basis of the X-ray crystal structure of WTcyt *c* helped in assigning the heme methyl resonances. Conversely, these same connectivities aided in determining the sequence-specific assignment of a number of amino acid side chains. In addition, a number of NOESY cross peaks between amino acid protons and heme substituents other than the methyls

were identified (Table 6.1). These were useful for verifying the heme proton assignments reported in Chapter 5, and also for determining the sequence-specific assignment of heme pocket residues.

Another important method used in performing the sequence-specific assignment is the identification of both dipolar and scalar sequential connectivities between backbone protons.[29] TOCSY and NOESY spectra acquired in H_2O allowed the assignment of backbone amide protons to particular spin systems, which were then linked together to learn the position of each amino acid in the sequence. Finally, the X-ray crystal structure of WTcyt *c*[30] was utilized to determine what interresidue NOESY cross peaks are expected from particular amino acids, assuming no major structural perturbations were induced in the examined area of the protein by the mutation. This was useful for areas of the protein in which NH-NH or Hα-NH sequential connectivities were not observed. This procedure, discussed in greater detail below, led to the assignment of 82% of all protons; 94% of the backbone protons, and 77% of the side-chain protons. The assignments are given in Table 6.2. Only the first two residues remain completely unidentified. These residues are also unidentified in the published assignments of oxidized and reduced *S. c.* cyt *c*.[31]

Amino Acid Side Chains

The resonances due to protons on amino acid side chains exhibit particular patterns in TOCSY and COSY maps according to their chemical structure and the relative chemical shifts among the protons.[29] The pattern expected differs depending on whether a COSY or TOCSY spectrum was acquired. In addition, varying the TOCSY mixing time changes the observed spin patterns, with longer mixing times allowing spin polarization transfer further along the side chain and giving a more complete spin pattern. Comparison of spin patterns obtained with different pulse sequences and mixing times,

Table 6.2

Proton NMR Assignments for S. c. CN-Fe(III)Ala80cyt c[a]

residue	NH (backbone)	α	β	others
Thr–5				
Glu–4				
Phe–3	8.68	4.39	2.64[1], 2.91[2]	δ 6.99; ε 7.24; ζ 7.08
Lys–2	6.63	3.84	1.46, 1.58	γ 1.19; δ 1.32; ε 2.89
Ala–1	7.68	3.64	1.27	
Gly 1	8.12	3.46[1], 3.65[2]		
Ser 2	9.44		3.87, 3.65	
Ala 3		4.02	1.49	
Lys 4	8.26	4.03	1.71, 1.83	ε 3.00
Lys 5	7.74	4.12	1.39, 1.75	γ 2.22
Gly 6	8.68	3.31[1], 4.03[2]		
Ala 7	7.94	2.29	1.18	
Thr 8	7.18	3.89	4.21	γ 1.24
Leu 9	7.93	3.88	1.90, 1.58	δ 0.72, 0.92
Phe 10	8.59	3.51	2.59[1], 2.87[2]	δ 6.87; ε 6.22, 7.76; ζ 6.40
Lys 11	8.15	4.22	1.78, 2.11	ε 3.06
Thr 12	7.93	4.32		γ 1.32
Arg 13	8.40	4.61	1.66[1], 1.43[2]	γ 1.77
Cys 14	7.59	–0.14	1.88[1]	
Leu 15	7.70	5.85	1.89[1], 2.30[2]	γ 2.24; δ 1.37; 1.96
Gln 16	9.50	4.47	2.01, 2.05	γ 2.34, 2.57; ε 6.61[1], 7.21[2]
Cys 17	9.03	6.09	2.50[2], 4.99[1]	
His 18	11.05	9.78	7.75[1], 11.94[2]	δ1 16.54, δ2 16.1; ε1 –3.4
Thr 19	10.52	6.14	5.47	γ 2.28
Val 20	9.08	4.83	1.88	γ 0.62, 0.67
Glu 21	9.10	4.43	1.98, 2.04	γ 2.19, 2.27
Lys 22	8.93	3.48	1.57[2], 1.62[1]	γ 0.79[2], 0.94[1]; δ 1.16[1]; ε 3.00[1]
Gly 23	9.53	3.85, 4.17		
Gly 24	8.38	3.83[2], 4.56[1]		
Pro 25		3.60	2.43, 2.01	
His 26	9.08	5.87	3.65[2], 3.71[1]	δ2 7.39, ε1 8.17
Lys 27	9.34	5.18	2.80[2], 2.89[1]	γ 2.44; δ 2.10[1], 2.53[2]; ε 3.62[2], 3.69[1]
Val 28	7.96	3.61	1.49	γ –0.49[1], 0.86[2]
Gly 29	9.97	3.88[1], 4.01[2]		
Pro 30		5.33	1.17, 2.44	γ 1.42, 2.73; δ 3.71
Asn 31	12.46	6.39	3.00[2], 3.35[1]	δ 8.89[1], 8.01[2]

Table 6.2 (cont.)

residue	NH	α	β	others
Leu32	9.64	4.54	1.23[2], 1.86[1]	γ 4.07; δ –0.87[1], –0.42[2]
His33	8.22	4.10	3.27, 3.34	δ2 7.35, ε1 8.22
Gly34	9.09	3.69[1], 3.80[2]		
Ile35	6.86	3.39	1.10	γCH_3 –0.51; γ –0.10,–0.78; δ –0.78
Phe36	7.57	3.92	2.70, 2.78	δ 6.81, ε 6.17, ζ 6.29
Gly37	8.64	3.71[1], 4.13[2]		
Arg38	7.84	4.63	1.84, 1.98	γ 1.71
Gln39	7.75	5.09	1.74, 2.15	
Ser40	8.43	4.42	3.44[2], 4.21[1]	
Gly41	7.71			
Gln42	7.85	4.54	1.77, 1.88	γ 2.14, 2.22; ε 6.70[2], 7.34[1]
Ala43	8.48	4.51	1.83	
Glu44	8.96	4.30	2.24	γ 2.45, 2.53
Gly45	7.25	3.94, 4.79		
Tyr46	7.17	4.21	1.24[1], 2.29[2]	δ 6.35, 6.69; ε 4.85, 5.86; η 9.29
Ser47	6.85	4.33	3.28[2], 3.41[1]	
Tyr48	7.89	4.50	2.60. ,3.29	δ 7.05, 7.20; ε 6.05, 7.12; η 9.42[b]
Thr49	9.28	4.22	4.49	γ 1.32, OH 12.0
Asp50	8.55	4.39	2.54[2], 2.62[1]	
Ala51	7.65	4.01	1.44	
Asn52	8.05	4.25	2.62[1], 2.73[2]	δ 6.34[2], 7.80[1]
Ile53	7.47	3.23	1.78	γCH_3 0.87; γ 1.02, 1.68; δ 0.94
Lys54	9.00	3.94	1.01, 1.85	γ 1.48, 1.58
Lys55	7.45	4.00	2.00, 1.96	γ 1.37; δ 2.78
Asn56	7.17	4.28	2.32, 2.97	δ 6.53; 7.43
Val57	7.26	3.76	0.98	γ –0.20[1], 0.41[2]
Leu58	8.11	4.13	1.29[1], 1.58[2]	γ 0.88; δ 0.67, 0.75
Trp59		4.72	2.04,[1] 3.26[2]	δ1 6.53; ε1 8.19; ε3 7.12; ζ2 6.73; ζ3 6.58; η2 6.36
Asp60	9.55		2.94, 2.71	
Glu61	9.99	3.34	1.35, 1.38	γ 0.89, 1.14
Asn62	8.13	4.54	2.87[1], 2.92[2]	δ2 7.11[2], 7.74[1]
Asn63	9.43	4.58	2.82, 3.13	δ2 6.92, 7.39
Met64	8.71	4.17	1.98[1], 1.53[2]	ε –0.41
Ser65	7.53	3.77	3.96, 4.25	
Glu66	7.77	4.15	2.17[1], 2.24[2]	γ 2.43[1], 2.63[2]
Tyr67	8.76	4.52	3.28[2], 3.52[1]	δ 7.25, 7.31; ε 7.59, 7.77; OH 19.3
Leu68	8.83	4.99	1.23[2], 1.73[1]	γ 1.35, δ –0.95, –0.54

Table 6.2 (cont.)

residue	NH	α	β	others
Thr69	7.97	4.37	4.73	γ 1.52
Asn70	7.68	5.54	3.54, 3.63	
Pro71		6.37	6.10^1, 8.82^2	γ 7.40, 9.47, δ 5.48^2, 5.80^1
Tml72[c]	9.62	5.08	2.15, 2.29	γ 2.56, 2.60; ε 3.73; $C(CH_3)_3$ 3.13
Lys73	8.24	4.66	2.26^1, 2.37^2	γ 1.81, 1.86; δ 2.07; ε 3.33
Tyr74	8.45	4.87	4.15^1, 4.37^2	δ 7.87, ε 6.98
Ile75	9.57	4.73	3.15	γCH_3 1.02; γ 1.80, 3.04; δ 1.70
Pro76		5.05	2.05, 2.55	γ 2.28; δ 3.80
Gly77	9.11	3.81, 4.41		
Thr78	9.61	4.64	5.08	γ 2.57
Lys79	8.92		2.41^1, 2.48^2	γ –0.22, 0.95; δ 1.34; ε 0.40
Ala80	7.18	–1.52	1.40	
Ala81	5.39	4.05	0.39	
Phe82	6.70		0.36	ζ 6.82 δ/ε 7.60
Gly83	7.54	4.26, 5.22		
Gly84		4.27, 5.11		
Leu85	9.04	4.87	1.62, 1.66	γ 1.51; δ 0.61^1, 0.80^2
Lys86	8.78	4.25	1.98, 2.02	γ 1.67, 1.71; δ 1.88, 1.90; ε 3.22
Lys87	8.93	4.51	1.46, 1.66	ε 3.05
Glu88	9.11			
Lys89	8.74	3.94	1.83, 1.88	γ 1.66, 1.71; δ 1.47, 1.51; ε 2.99
Asp90	6.40	4.20	2.43^1, 2.88^2	
Arg91	7.52	3.53	1.86^1, 2.80^2	γ 1.04; δ 2.37
Asn92	8.59	4.51	2.71	
Asp93	8.52	4.11	2.51^2, 2.58^1	
Leu94	8.31	3.82	1.51, 1.44	γ 1.01; δ 0.17, 0.31
Ile95	8.75	2.92	1.77	γCH_3 0.17; γ 0.23; δ 0.60
Thr 96	7.90	3.68	4.29	γ 1.14
Tyr97	7.60	4.02	2.83^2, 3.51^1	δ 6.61, 6.44; ε 7.00, 7.23
Leu98	8.93	3.00	0.98, 1.05	γ 1.11; δ –0.63, –0.23
Lys99	8.46	3.26	1.12^2, 1.38^1	γ 0.44^1, 0.70^2; δ 2.44^2, 0.64^1
Lys100	6.46	4.12	1.72, 1.75	γ 1.37; δ 1.60, ε 2.95
Ala101	8.22	3.94	0.53	
Ser102	7.61	4.30	3.28	
Glu103	6.83	4.00	1.94^1, 2.01^2	γ 2.27, 2.32

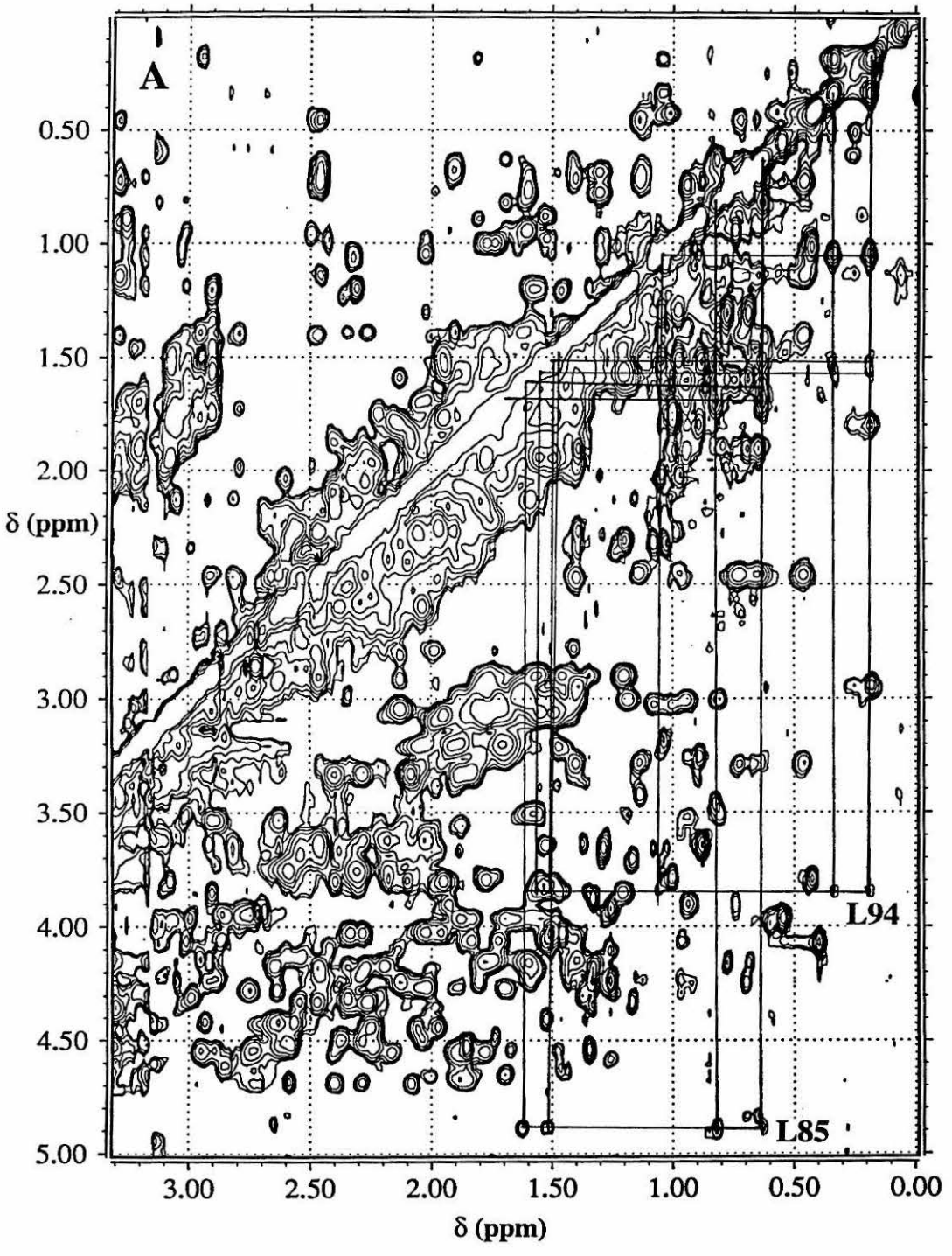
[a]Shifts measured at 303 K unless noted otherwise. [1,2] refer to stereospecifc assignments.
[b]Shift measured at 295 K.
[c]trimethyllysine.

therefore, is valuable in identifying particular protons (*i.e.* Hβ, Hγ) within a side chain. Examples of some identified spin systems are presented in Figures 6.2 and 6.3. In Figure 6.2A, the spin system of two leucine side chains (identified by the sequence-specific assignment as Leu85 and Leu94, see below) are indicated in a TOCSY spectrum obtained with a long (90 ms) spin-lock time. TOCSY data collected with a long spin-lock time greatly facilitated the identification of long side chains, as it allowed the observation of coupling from the Hα to most other side-chain protons. This is valuable because the cross peaks between, for example, the Hα and the δ-methyl groups of leucine residues appear in a relatively uncrowded region of the spectrum (Figure 6.2A).

The 90-ms TOCSY map was valuable for identifying many of the 16 Lys side chains. In the published assignment of *S. c.* Fe(III)Thr102cyt *c*, no assignments are reported for protons other than the NH, Hα, and Hβ's for any Lys.[31] For *S. c.* CN-Fe(III)Ala80, at least three side-chain protons were identified for every Lys in the sequence; in many cases six or seven side-chain protons were assigned (Table 6.2). In Figure 6.2B, the spin system of a Lys (Lys 27) side chain is indicated. Connectivities were observed all down the side chain, from the Hα to the β, γ, δ, and ε protons. As with the Leu side chains, it was helpful to observe coupling to the Hα for all of the protons as this gives cross peaks in a relatively uncrowded spectral region. The intensities of these peaks can be qualitatively related to the number of bonds between the two protons; in Figure 6.2B, the Hα/Hε cross peaks are the weakest, while the Hα/Hβ cross peaks are relatively intense.

The aromatic region of a TOCSY (90-ms spin-lock time) map is shown in Figure 6.3. The pattern of Trp59, the only Trp present in the sequence, is clearly evident, as are the spin patterns of the rings of Phe-3, Phe10, Phe36, Tyr46, Tyr48, Tyr74, Phe82, and Tyr97. The nonexchangeable protons of His26 and His33 were also identified, thus accounting for all aromatic residues except for Tyr67, which is observable only in the

Figure 6.2 Aliphatic region of the 600 MHz ^{1}H NMR TPPI TOCSY (90-ms spin-lock time) spectrum of *S. c.* CN-Fe(III)Ala80cyt *c* in 50 mM NaP_i (uncorrected pH 7.0) in D_2O. The spin systems of the side chains of (A) Leu85 and Leu94 and (B) Lys27 are indicated.


A
0.50
1.00
1.50
2.00
δ (ppm)
2.50
3.00
3.50
4.00
4.50
5.00
L94
L85
3.00
2.50
2.00
1.50
1.00
0.50
0.00
δ (ppm)

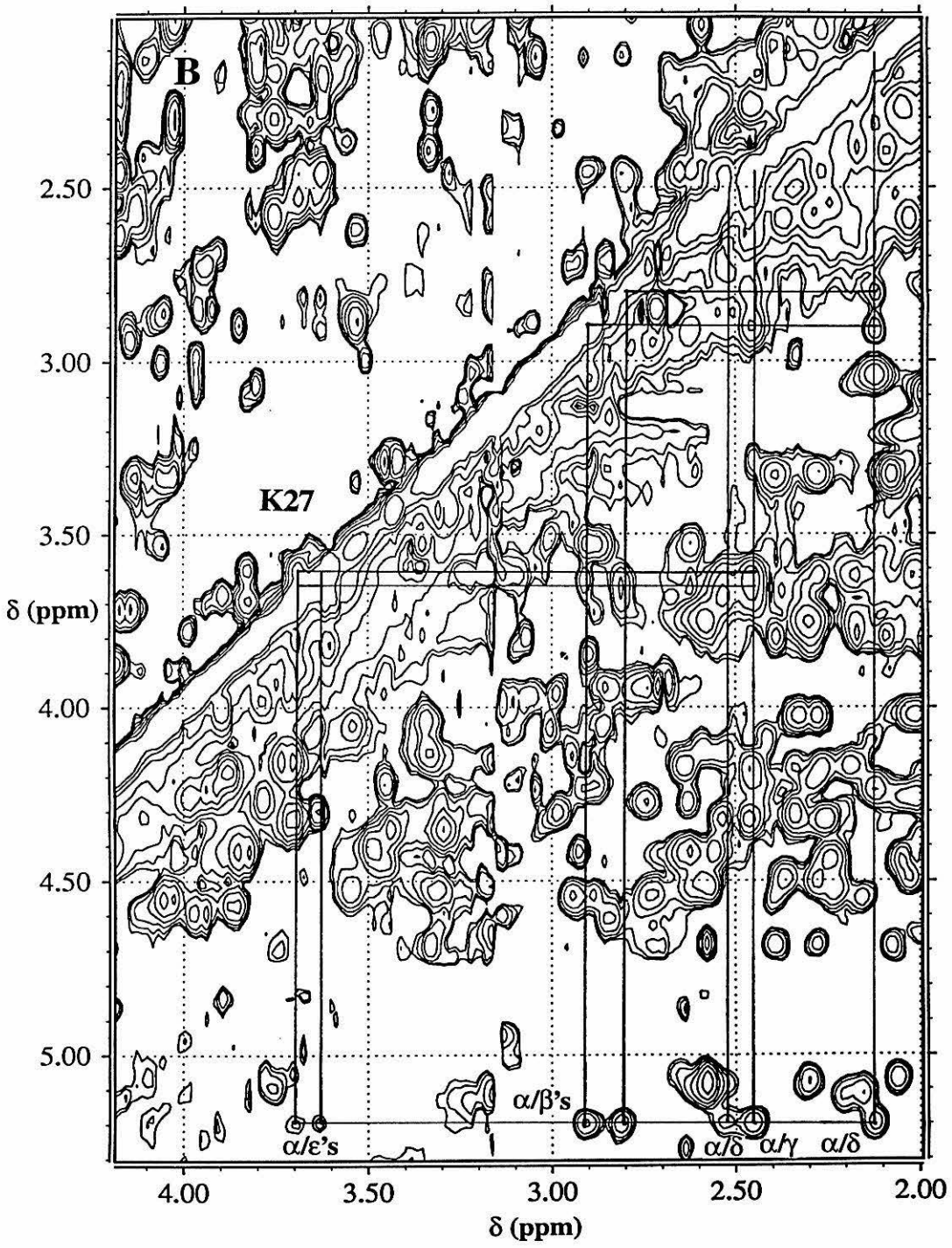
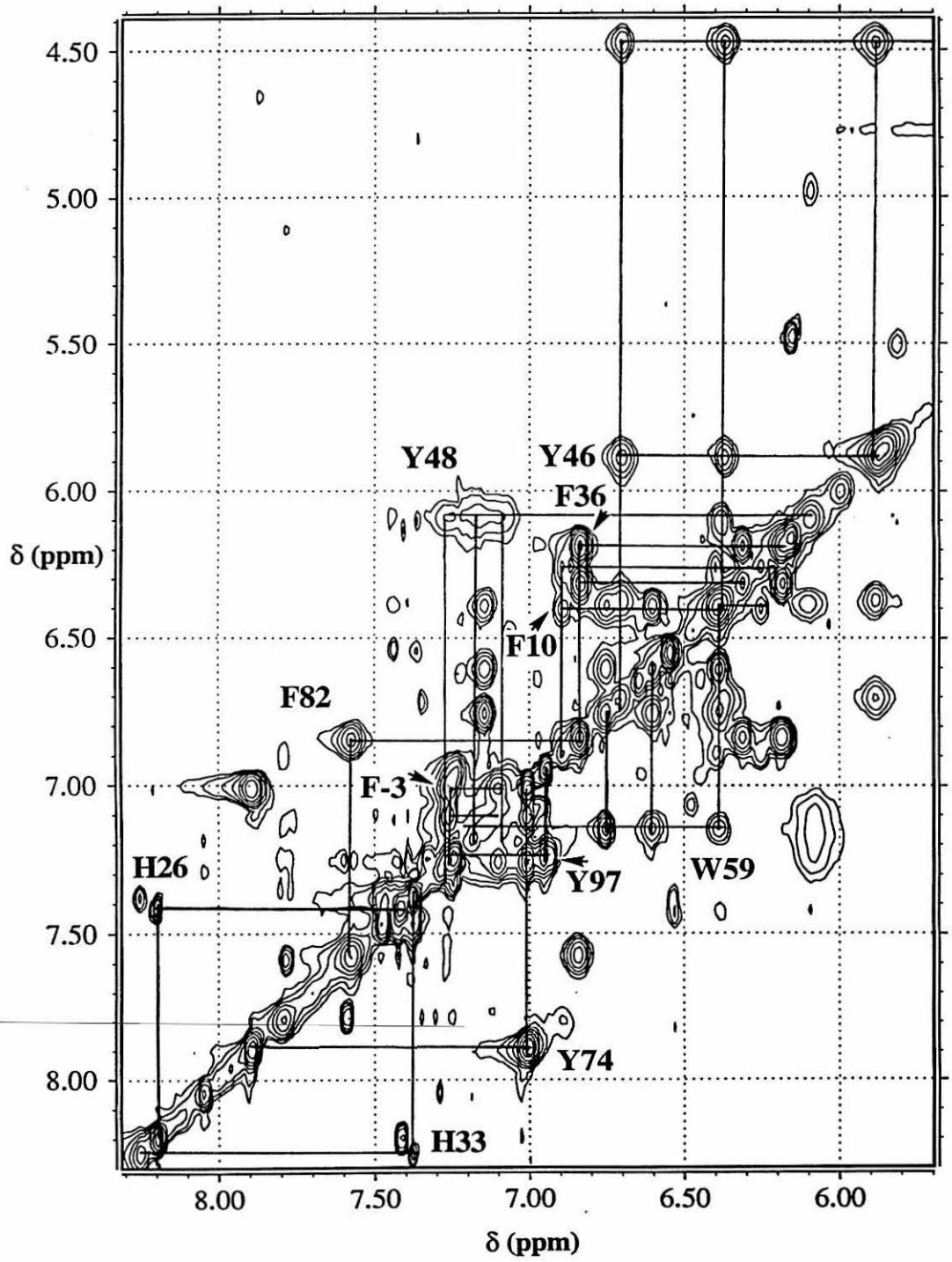

B
K27
2.50
3.00
3.50
4.00
4.50
5.00
δ (ppm)
α/ε's
α/β's
α/δ
α/γ
α/δ
4.00
3.50
3.00
2.50
2.00
δ (ppm)

Figure 6.3 Aromatic region of the 600 MHz ^{1}H NMR TPPI TOCSY (90-ms spin-lock time) spectrum of *S. c.* CN-Fe(III)Ala80cyt *c* in 50 mM NaP_i (uncorrected pH 7.0) in D_2O. The spin systems of all of the aromatic amino acids in the sequence, except for Tyr67 which is observed only at shorter spin-lock times, are indicated.


4.50
5.00
5.50
6.00
6.50
7.00
7.50
8.00
δ (ppm)
8.00
7.50
7.00
6.50
6.00
δ (ppm)
Y48
Y46
F36
F10
F82
F-3
H26
Y97
W59
Y74
H33

TOCSY spectrum collected with a spin-lock time of 30 ms (Figure 5.10). Tyr67 can only be seen with the shorter spin-lock time because of the vicinity of the Tyr67 ring to the paramagnetic center which results in a large reduction of the T_1 and T_2 values of its nuclei. The assignment of this residue has not previously been reported. The inability to detect this group in other cyts *c* was attributed to the flipping of the Tyr67 ring at a rate that is intermediate on the NMR timescale,[32,33] which is not the case in this system. The slow Tyr67 hydroxyl proton exchange rate as well as its large isotropic shift in *S. c.* CN-Fe(III)Ala80 (Chapter 5) suggests that this proton is hydrogen-bonded with the heme-CN unit. This may slow the ring-flipping rate relative to that observed in WTcyt *c*.

The analysis of the aromatic side-chain assignments for the Tyr and Phe residues provides information on the mobility of these rings (summarized in Figure 6.4). As reported for Thr102cyt *c*,[31] the observation of only one δ/ε scalar connectivity for Phe-3, Phe36, and Tyr74 (Figure 6.3) indicates that their rings are flipping rapidly. For Phe82, only two resonances were detected at all temperatures investigated. The assignment of the ring protons is based on NOESY connectivities observed from these resonances. Only two resonances for the Phe82 ring were observed also in Thr102cyt *c*; these data suggest that the ring is flipping rapidly and that two of the chemical shifts are nearly degenerate, or that one proton is not observed due to paramagnetic or other effects. The observation of more than one set of δ/ε resonances for Phe10, Tyr46, Tyr48, Tyr67, Tyr97 indicates that they are flipping slowly on the NMR timescale. This is in accordance with the reported results on other cytochromes *c*.[31-33]

In the aromatic region, the spin system of a proline residue (Pro71) was identified (Figure 6.5). The unusual shift range (Table 6.3) can be interpreted in terms of a significant pseudocontact contribution to the shift of these resonances, consistent with the essentially axial position of this residue with respect to the heme moiety (see equation 4.4) and the relatively short distances of its protons from the heme iron (within 6–9Å).[8,9]

Figure 6.4 Dynamics of aromatic rings in *S. c.* CN-Fe(III)Ala80cyt *c* compared to *S. c.* Fe(III)Thr102cyt *c*[31] as observed by ^{1}H NMR. The side chains of the Phe and Tyr residues in the cyt *c* sequence are shown along with a (one-thread) ribbon trace of the cyt *c* backbone from the X-ray crystal structure (PDB identifier 1CYC). Next to the residue names, the two letters indicate the rate (S = slow, R = rapid) on the NMR timescale at which that side chain flips. The rate for *S. c.* Fe(III)Thr102cyt *c* is given first, followed by that observed for *S. c.* CN-Fe(III)Ala80cyt *c*. For Tyr67, "_" indicates that it was not assigned for Thr102cyt *c*; this may be due a flipping rate that is intermediate on the NMR time scale.

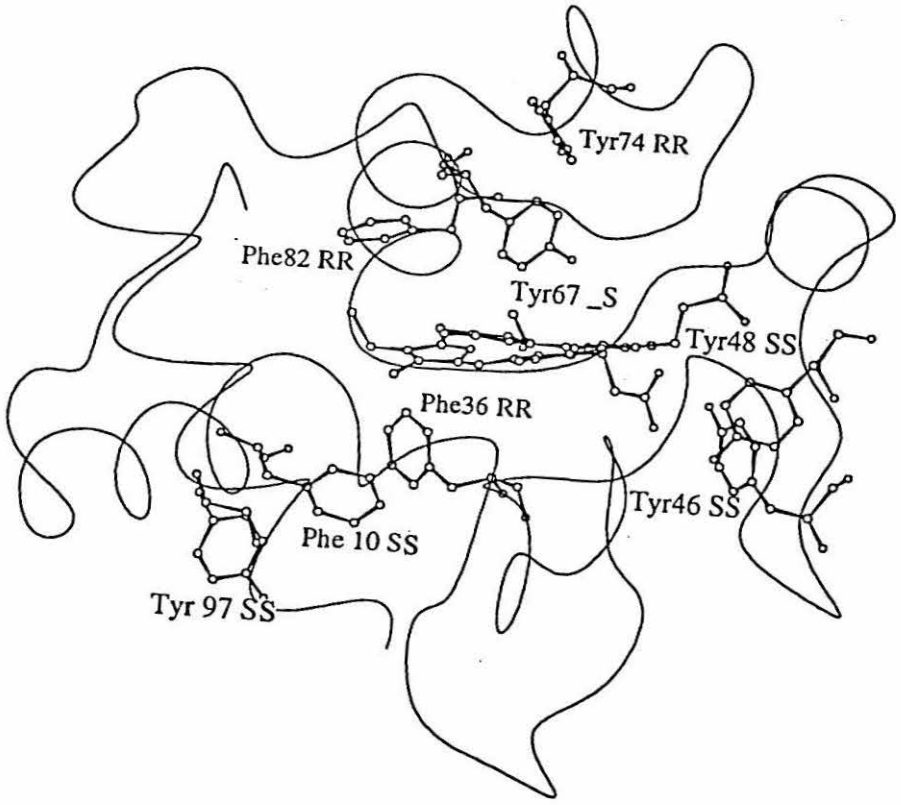
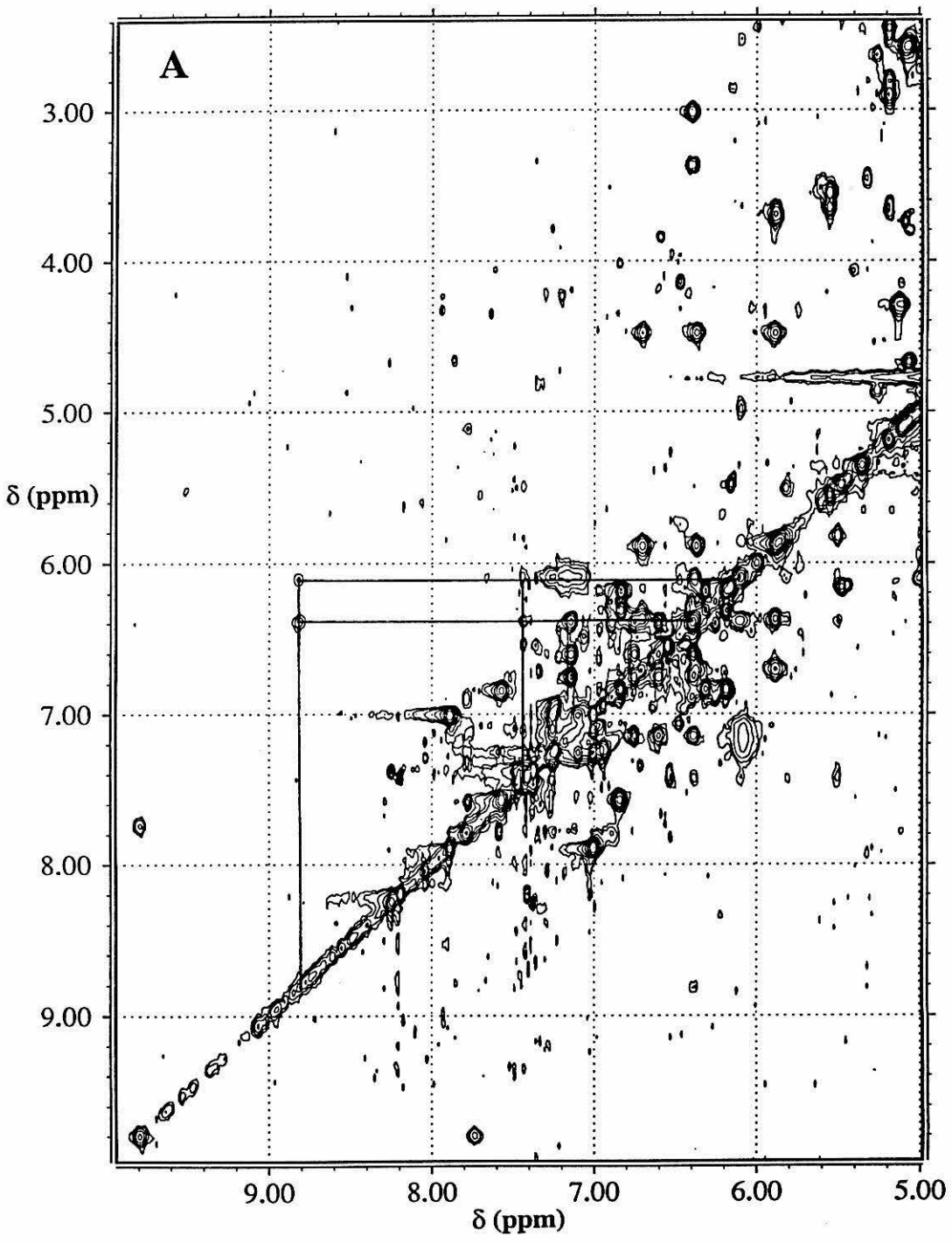
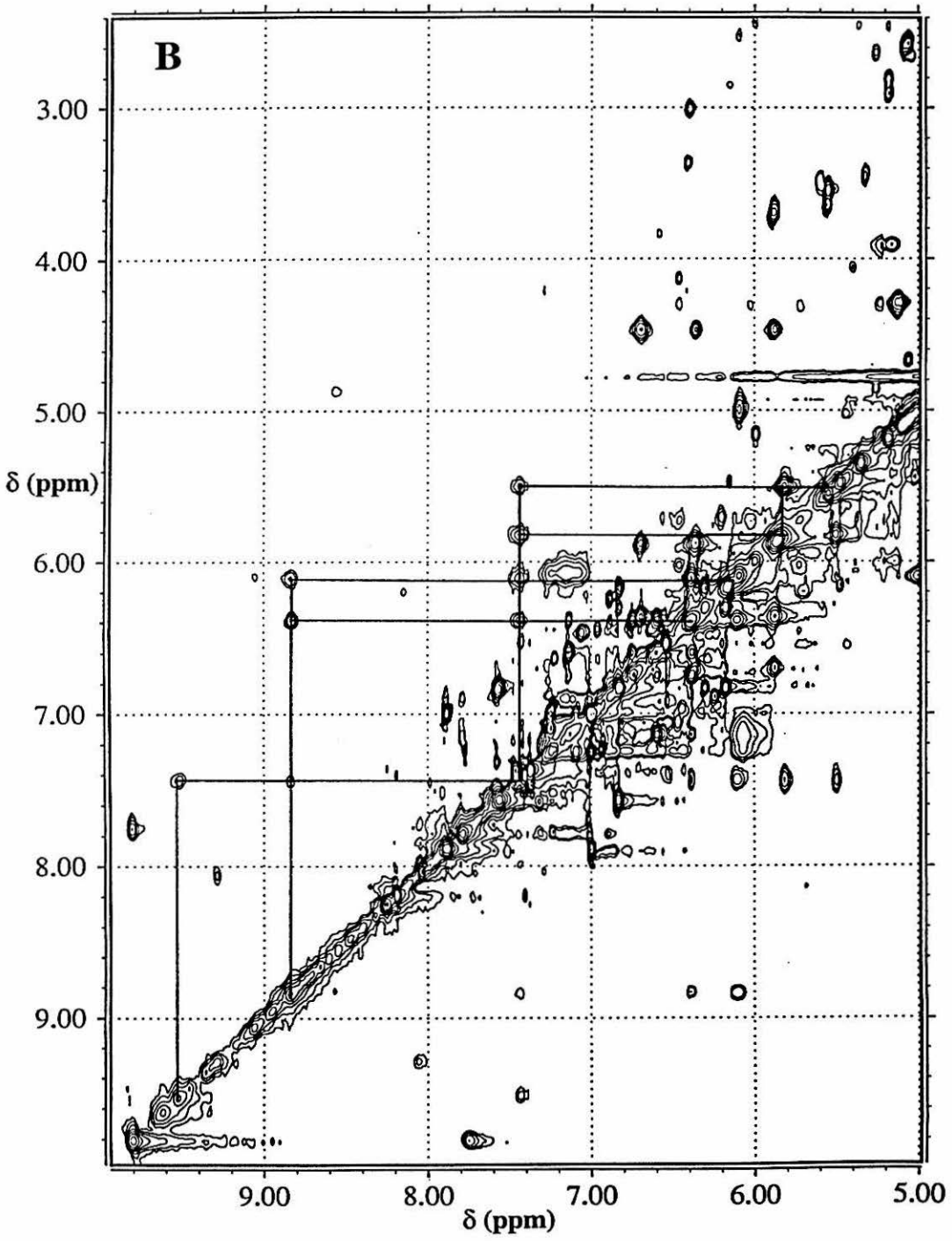

Tyr74 RR
Phe82 RR
Tyr67 _S
Tyr48 SS
Phe36 RR
Tyr46 SS
Phe 10 SS
Tyr 97 SS

Figure 6.5 600 MHz ^{1}H NMR spectra (50 mM NaP_i, uncorrected pH 7.0, D_2O, 303 K) with the Pro71 protons indicates. (A) TPPI TOCSY spectrum taken with a spin-lock time of 90 ms. Because the Pro71 protons relax rapidly, the cross peaks are weak and only four protons are detected. (B) TPPI TOCSY spectrum taken with a 30-ms spin-lock time. In this spectrum, cross peaks from all 7 of the Pro71 resonances are visible. (C) TPPI NOESY spectrum taken with a short (15-ms) mixing time. All 7 of the Pro71 resonances can be identified, giving broad, intense cross peaks.


A
3.00
4.00
5.00
6.00
7.00
8.00
9.00
δ (ppm)
9.00
8.00
7.00
6.00
5.00
δ (ppm)


B
3.00
4.00
5.00
6.00
7.00
8.00
9.00
δ (ppm)
9.00
8.00
7.00
6.00
5.00
δ (ppm)

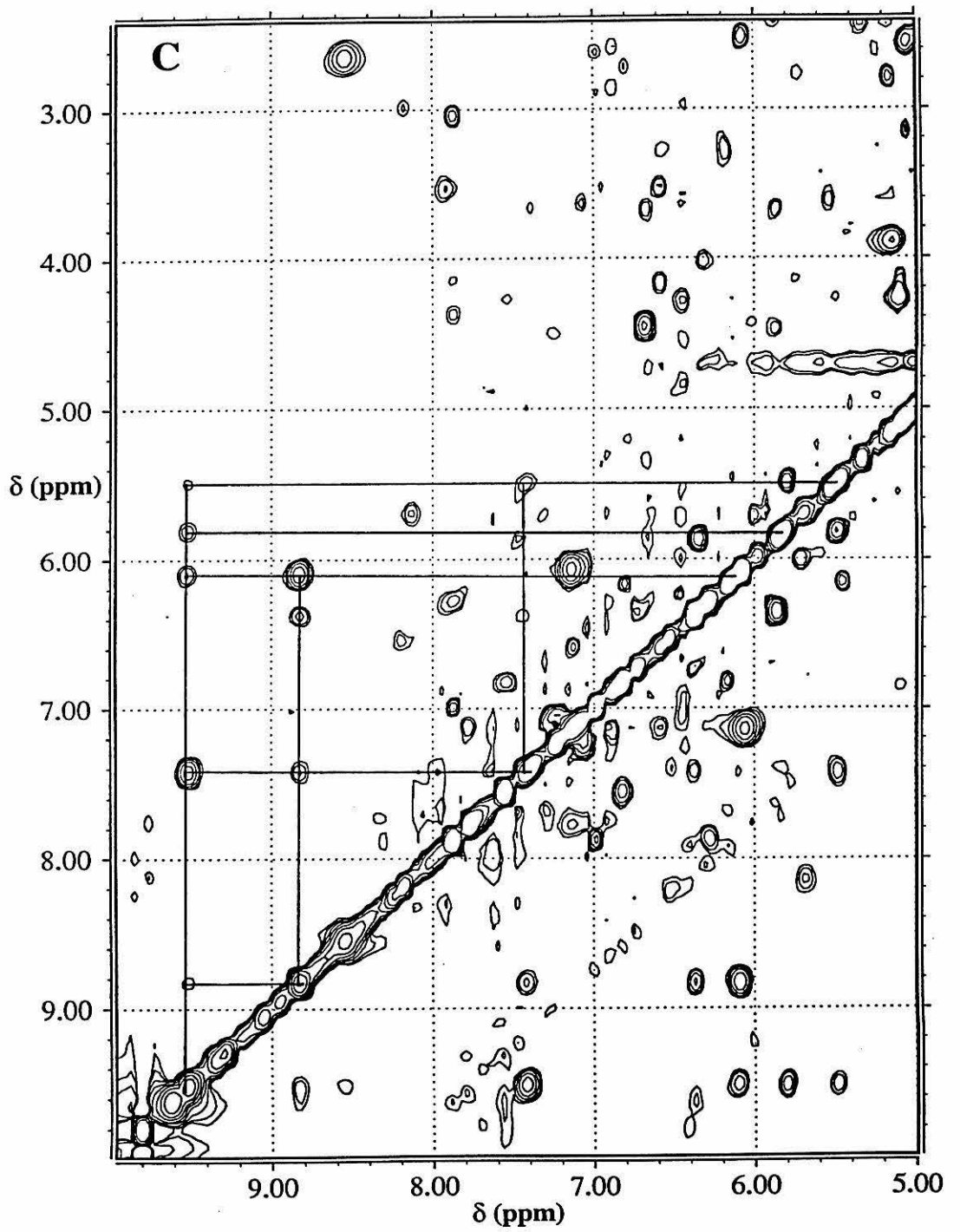

C
3.00
4.00
5.00
6.00
7.00
8.00
9.00
δ (ppm)
9.00
8.00
7.00
6.00
5.00
δ (ppm)

Table 6.3

^1H NMR chemical shifts of Pro71 in S. c. CN-Fe(III)Ala80cyt c compared to a random-coil proline

residue	Hα	Hβ	Hβ	Hγ	Hγ	Hδ	Hδ
typical Pro[a]	4.44	2.28	2.02	2.03	2.03	3.65	3.68
Pro71	6.37	8.82	6.10	7.40	9.47	5.48	5.80

[a]from reference 29; data for *trans*-Pro in a random-coil polypeptide

The assignment of the Pro71 resonances is supported by interresidue NOESY connectivities to nearby residues including Tyr67. 1D NOE connectivities were also detected between one Hγ proton and the two Hβ protons of Pro71 and the OH of Tyr67. The position of this residue also results in a substantial increase in the relaxation rates of its nuclei, as is apparent by comparing the Pro71 spin pattern in TOCSY spectra taken with different spin-lock times. In Figure 6.5A, what can be seen of the spin pattern of Pro71 in a TOCSY spectrum taken with a 90-ms spin-lock time is shown. Decreasing the spin-lock time to 30 ms (Figure 6.5B) results in the appearance of many more of the Pro71 cross peaks. Intraresidue NOESY peaks for Pro71 are very prominent in the NOESY spectrum acquired with a short (15-ms) mixing time (Figure 6.5C).

Unusual shift values were also observed for the resonances of the two heme-bound cysteines (Cys14, Cys17) and the heme iron axial ligand His18, as well as Thr19, Asn31, Leu32, Ile35, Lys79, Ala80, and Ala81. All of these residues are within a 9-Å sphere from the heme iron.

Exchangeable side chain protons were identified for Gln16, Asn31, Gln42, Tyr46, Tyr48, Thr49, Asn52, Asn56, Trp59, Asn62, Asn63, and Tyr67. The observation of a Tyr hydroxyl proton suggests that it is sequestered from solvent, most likely by hydrogen bonding.

Backbone Protons

Comparative analysis of the dipolar and scalar connectivities between backbone protons (Figure 6.6) observed in the amide and fingerprint regions of the 2D NOESY and COSY spectra allowed the sequence-specific assignment of the backbone protons of 98 amino acids (Table 6.2, Figure 6.7). TOCSY spectra in water were used to identify scalar connectivities between the side-chain (typically Hβ and Hγ) and backbone NH protons of several residues. In many parts of the sequence, NH-NH (Figure 6.8) and NH-Hα (Figure 6.9) connectivities were observed. It was convenient to start the backbone proton assignment from His18, as its assignment is unequivocal and its NH and Hα resonances well-resolved (Figure 5.13). NH-NH connectivities were observed from His18 to Cys17, and on to Lys5, as shown in Figures 6.7A, B. Throughout the sequence, sequential NH-NH connectivities for stretches longer than two residues were observed for positions -3 to -1, 4–14, 15–21, 27–29, 31–33, 36–38, 50–56, 61–67, 68–70, 72–75, 88–92, and 93 to the C-terminus. Residues 25, 30, 72, and 76 are prolines, which do not contain the amide NH, and therefore sequential NH connectivities cannot be observed for them. Possible NH-NH connectivities at positions 14, 26, 38, 41, 56, 67, 87, and 92 cannot be observed because of nearly degenerate amide proton chemical shifts at positions i and $i + 1$. Connectivities involving the NH of residues -5, -4, 3, 59, and 84 are unobservable because the corresponding resonances have not been identified. Besides these unassigned amides, all other short- and medium-range backbone NOEs expected on the basis of the known cytochrome fold and the final structure were observed, with three exceptions. The distances between the amide protons of residues 1 and 2 and 43 and 44 are both ~ 3.4 Å; however, the corresponding connectivities are expected at the same position on the NOESY map as those between residues 62 and 63 and between residues 98 and 99, respectively. The latter two cross peaks are much more intense as they correspond to shorter distances (~ 2.6–2.7 Å), therefore precluding the observation of the former cross peaks. The expected NH-NH connectivity between residues 44 and 45 has

Figure 6.6 Illustration of a peptide backbone and the dipolar (NOESY) and scalar (COSY) connectivities observed among its protons. COSY correlates nuclei that are separated by not more than three bonds; as a result, in the fingerprint (NH/Hα) region of the ^{1}H NMR COSY spectrum acquired in water, each peak corresponds to the NH/Hα connectivity within one amino acid. NOESY correlates protons close to each other in space (generally < 4.5 Å) and is often observed between sequential NH protons as well as the Hα of a residue and the NH of the following residue.

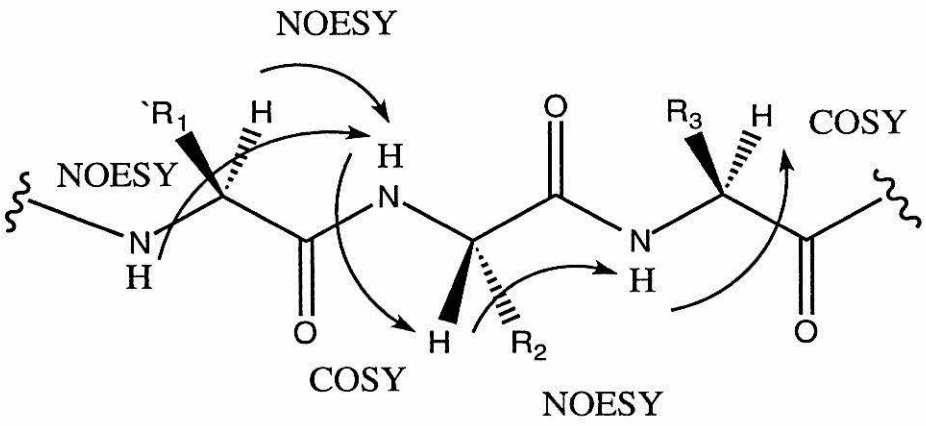

NOESY
R1
H
NOESY
N
H
H
N
O
H
R2
O
COSY
NOESY
R3
H
COSY
N
H
O

Figure 6.7 Schematic representation of the sequential and medium-range connectivities involving NH, Hα, and Hβ protons in *S. c.* CN-Fe(III)Ala80cyt *c*. For the sequential connectivities, the thickness of the bar indicates the NOE intensities. The medium-range NOEs are identified by lines connecting the two correlated residues. In the first line, NH nuclei found to exchange slowly with D_2O solution are indicated.

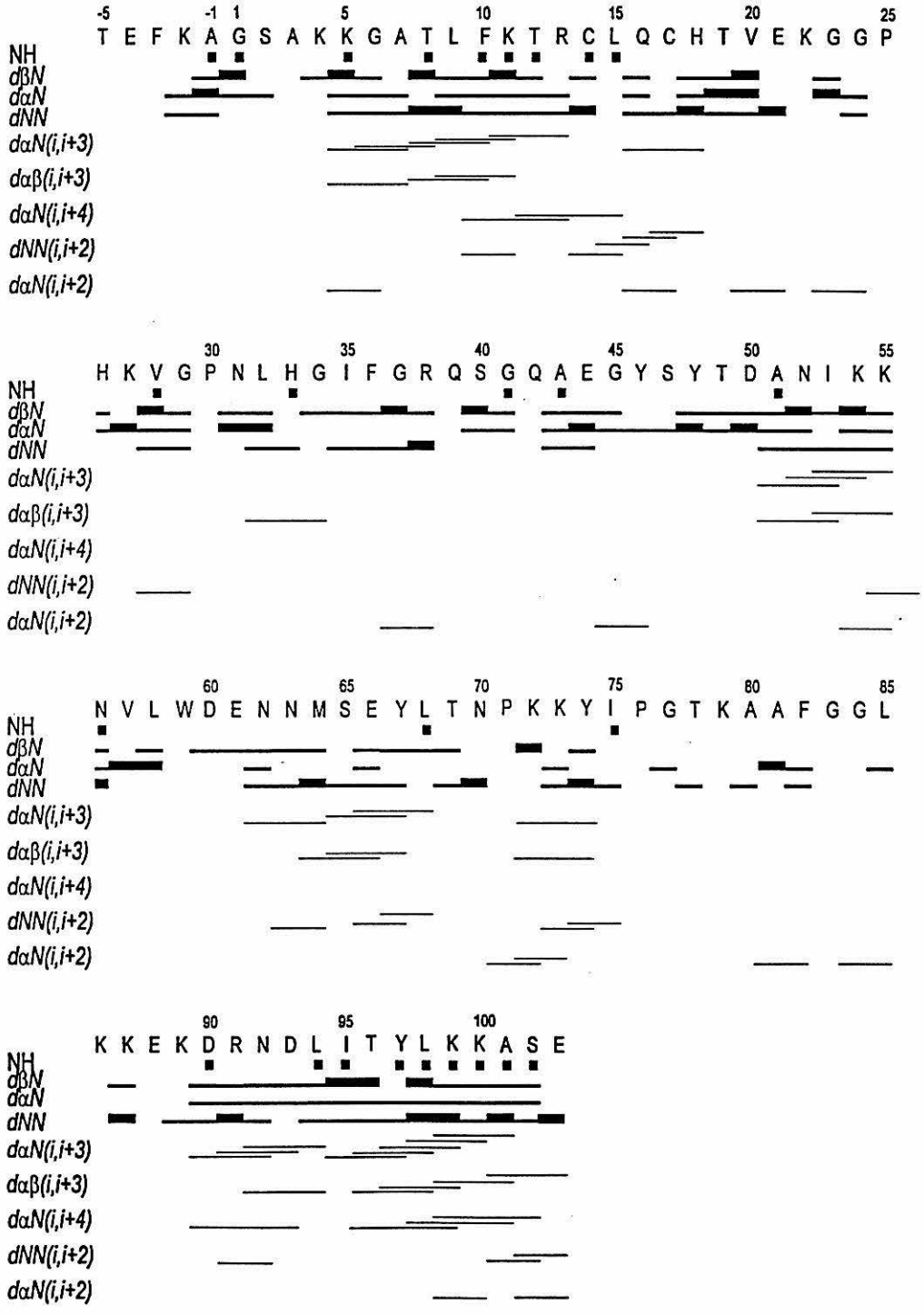
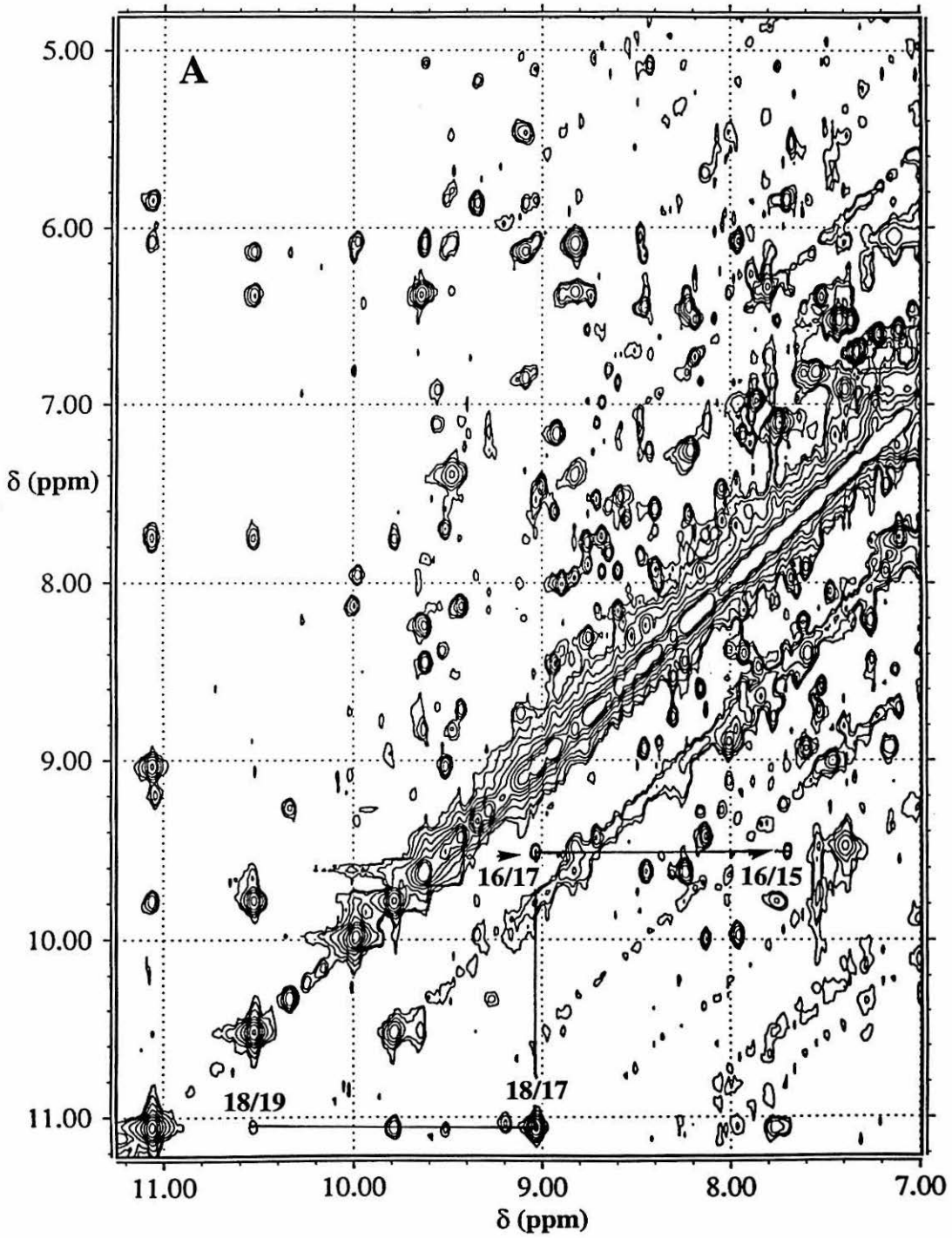

-5 -1 1 5 10 15 20 25
T E F K A G S A K K G A T L F K T R C L Q C H T V E K G G P
NH
dβN
dαN
dNN
dαN(i,i+3)
dαβ(i,i+3)
dαN(i,i+4)
dNN(i,i+2)
dαN(i,i+2)
30 35 40 45 50 55
H K V G P N L H G I F G R Q S G Q A E G Y S Y T D A N I K K
NH
dβN
dαN
dNN
dαN(i,i+3)
dαβ(i,i+3)
dαN(i,i+4)
dNN(i,i+2)
dαN(i,i+2)
60 65 70 75 80 85
N V L W D E N N M S E Y L T N P K K Y I P G T K A A F G G L
NH
dβN
dαN
dNN
dαN(i,i+3)
dαβ(i,i+3)
dαN(i,i+4)
dNN(i,i+2)
dαN(i,i+2)
90 95 100
K K E K D R N D L I T Y L K K A S E
NH
dβN
dαN
dNN
dαN(i,i+3)
dαβ(i,i+3)
dαN(i,i+4)
dNN(i,i+2)
dαN(i,i+2)

Figure 6.8 600 MHz ^{1}H NMR TPPI NOESY spectra (50 mM NaP_i (pH 7.0), H_2O, 303 K) of *S. c.* CN-Fe(III)Ala80cyt *c*. Sequential dipolar connectivities between the backbone amide protons of residues (A) 18–15 and (B) 15–5 are indicated. The amide protons of residues 7 and 9 have degenerate shifts at this temperature, causing overlap of the corresponding peaks.


A
5.00
6.00
7.00
8.00
9.00
10.00
11.00
δ (ppm)
16/17
16/15
18/19
18/17
11.00
10.00
9.00
8.00
7.00
δ (ppm)

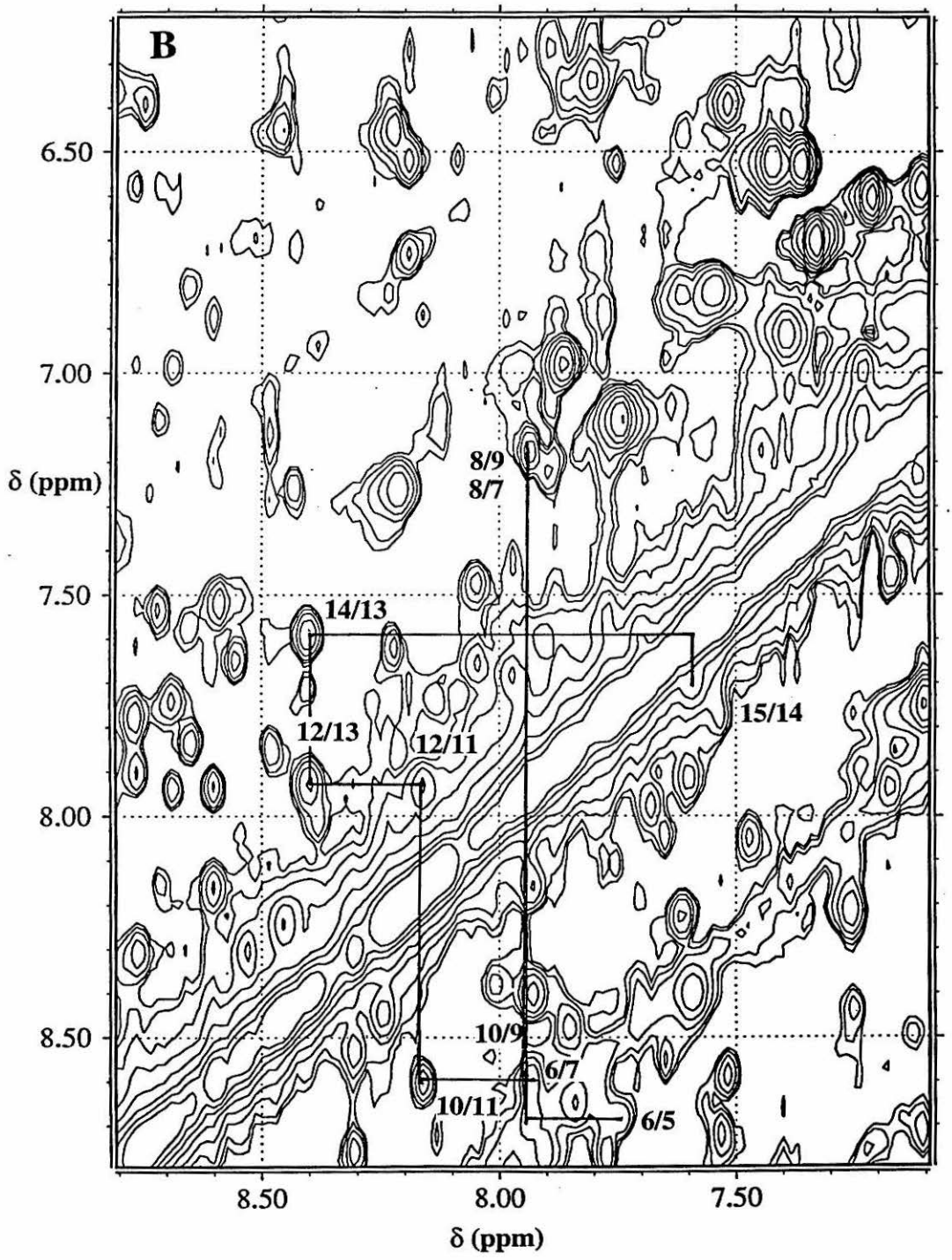
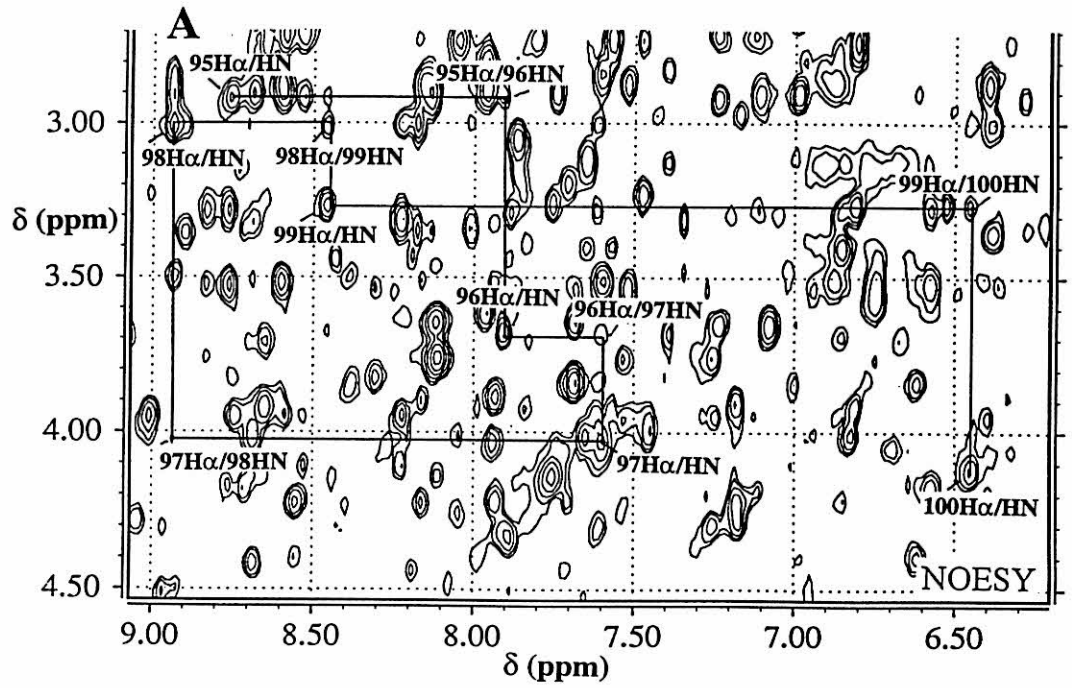
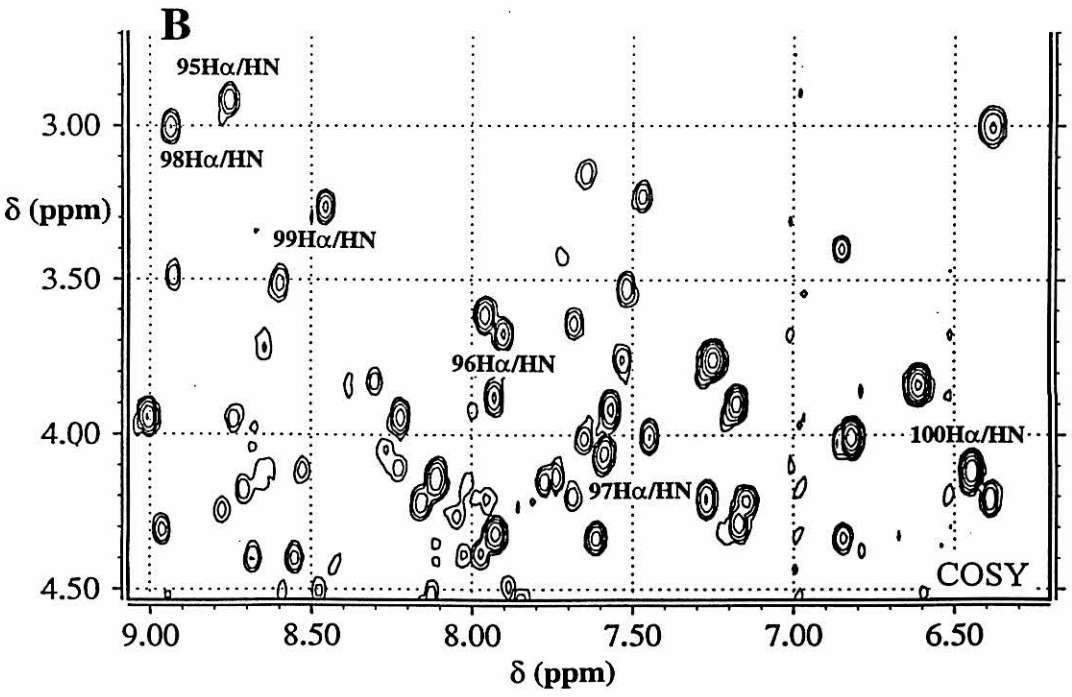

B
6.50
7.00
7.50
8.00
8.50
δ (ppm)
8.50
8.00
7.50
δ (ppm)
8/9
8/7
14/13
12/13
12/11
15/14
10/9
6/7
10/11
6/5

Figure 6.9 Fingerprint region of the 600 MHz ^{1}H NMR (A) TPPI NOESY and (B) COSY maps (50 mM NaP_i (pH 7.0), H_2O, 303 K) of *S. c.* CN-Fe(III)Ala80cyt *c*. (A) Sequential and intraresidue dipolar and (B) intraresidue scalar connectivities involving backbone protons for amino acids 95–100 are indicated.


A
95Hα/HN
95Hα/96HN
98Hα/HN
98Hα/99HN
99Hα/100HN
99Hα/HN
96Hα/HN
96Hα/97HN
97Hα/98HN
97Hα/HN
100Hα/HN
NOESY
3.00
3.50
4.00
4.50
δ (ppm)
9.00
8.50
8.00
7.50
7.00
6.50
δ (ppm)


B
95Hα/HN
98Hα/HN
99Hα/HN
96Hα/HN
100Hα/HN
97Hα/HN
COSY
3.00
3.50
4.00
4.50
δ (ppm)
9.00
8.50
8.00
7.50
7.00
6.50
δ (ppm)

not been detected. The low intensity of the position 44 and 45 intraresidue NH-Hα cross peaks suggests that these amide protons, which are positioned on the surface of the protein, are exchanging very rapidly with solvent, thus preventing the detection of the NH-NH cross peak.

An additional means used to assign the resonances due to backbone and corresponding side-chain protons is the identification of backbone NH-Hα connectivities.[29] This procedure, illustrated in Figure 6.6, begins with the identification of a COSY cross peak between a NH and Hα of one residue (*i*) assigned to a particular residue or residue type. Because COSY connectivities are not observed for protons separated by more than three bonds, this peak necessarily represents an intraresidue connectivity. Then, looking along the chemical shift of the Hα in a NOESY spectrum obtained in H_2O, it is possible to identify the NH of the next residue in the sequence (i + 1). Staying on the chemical shift of this NH, one should find an intraresidue cross peak between the NH and the Hα of residue i + 1. The COSY and NOESY data with the connectivities for residues 95–100 are shown in Figure 6.9. These types of connectivities for stretches longer than 2 residues were identified for residues -3–2, 4–7, 8–13, 17–20, 22–24, 26–29, 30–32, 39–41, 42–52, 53–58, 80–82, and 89–102 (Figure 6.7). As the sequence-specific assignment continued by a combination of these methods, assignments were checked for self-consistency by identifying long-range interresidue NOESY cross-peaks and determining whether these are reasonable based on the X-ray crystal structure of WTcyt *c*.[30]

Information on the Secondary Structure

Once the backbone and side-chain proton assignments were obtained, the NOESY spectra were reexamined for medium-range sequential NOEs. Backbone (i, i + 3) NOEs are indicative of α-helical structure.[29] As reported in Figure 6.7, Hα-Hβ (i, i + 3) connectivities were observed for residues 4, 7, 8, 31, 50, 52, 63, 64, 71, 91, 95, 96, 98,

and 100. Hα-NH (i, i + 3) connectivities, also indicative of a helix, were observed for residues 9, 11, 89, 95, 97, and 98. These data indicate the presence of α-helical motifs for residues 4–15, 50–55, 61–68, 70–75, and 89–103. Backbone (i, i + 2) NOEs are indicators of 3_{10} helical elements.[29] Interestingly, NOESY connectivities of this kind are observed for several of the residues in these helices, suggesting distortion of all of the helices although to different extents. NH-NH (i, i + 2) NOEs were also observed for the 13–18 region, which contains the two thioether linkages to the heme (Cys14, Cys17) and the heme axial ligand (His18). Analysis of the data reported in Figure 6.7, therefore, establishes that the five helical elements present in all of the characterized class I eukaryotic cytochromes *c* are retained in the Ala80 variant. As already reported for *S. c.* Thr102 iso-1-cyt *c*,[31] sequential NH-NH, Hα-NH, and Hβ-NH connectivities were also observed for the Phe-3–Ser2 region. This can be taken as evidence that the N-terminal region possesses inherent structure in solution, although not of helical type.

Solution Structure Determination

NOE Constraints

A total of 1834 experimental NOESY constraints were obtained using the program CALIBA (see Appendix B). Additional distance constraints have been included from 1D NOE difference spectra obtained by saturating the hyperfine-shifted signals of the ring protons of His18 and the Tyr67 OH. The intensities of these connectivities were measured and scaled to those obtained from 2D measurements. Out of a total of 1842 experimental distance constraints, 1426 have been used in DG and consequently in REM and RMD calculations. A total of 416 remaining constraints were found to be irrelevant for the calculation (*i. e.*, either no conformation of the polypeptide chain could violate them or they correspond to fixed proton-proton distances). For protons without stereospecific assignments, pseudoatoms were used. The number of constraints transformed into upper distance limits is presented in Table 6.4 for each calibration class.

Table 6.4

Constraints constituting the upper distance limit file for each calibration class

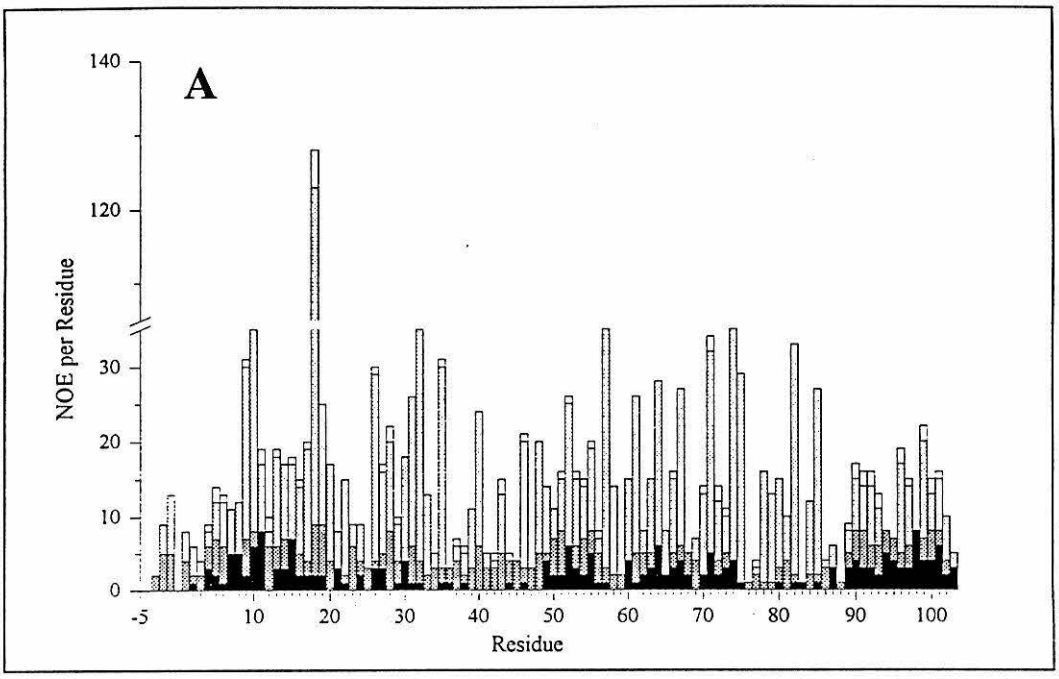
class	definition	number of constraints (methyl peaks) experimental	meaningful
1	intraresidue (except NH, Hα, Hβ)	504 (142)	274 (87)
2	sequential and intraresidue NH, Hα, Hβ	601 (23)	418 (14)
3	medium range	106 (3)	106 (3)
4	long range backbone	22	22
5	long range	609 (269)	606 (296)
total number of upper distance constraints		1842	1426

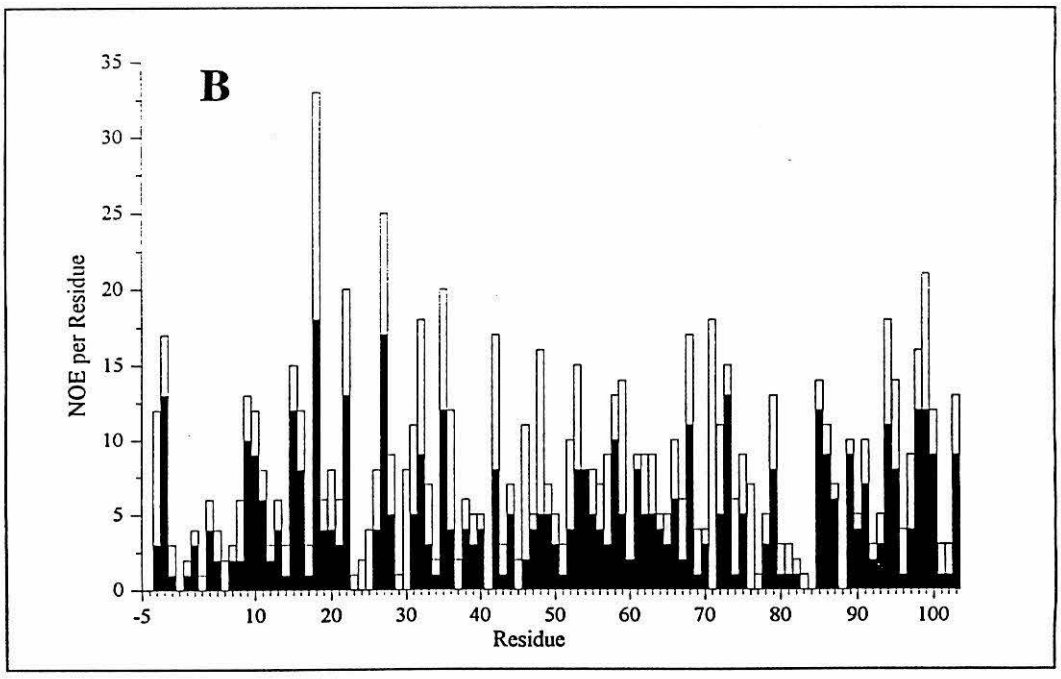
The calibration was in agreement with the volume-to-distance correlation when the intensities were assumed to be inversely proportional to the power of 4 for the corresponding upper distance limits in the case of protons in classes 1, 3, 4, and 5 and for methyls in classes 1, 2, and 5; to the power of 6 for protons in class 2; and to the power of 5 for methyls in class 3. Exponents lower than 6 have commonly been found to give better calibration curves as they empirically take into account spin diffusion and possible contributions to the reorientational correlation time.[13-15] The number of NOE constraints per residue is depicted graphically in Figure 6.10. With this number of distance constraints (which corresponds to 17.0 NOEs per residue as input for the DG calculation and to 13.1 accepted experimental constraints per residue), a good-quality solution structure is expected.[34]

Definition of the Prosthetic Group for Distance Geometry Calculations

The cyanide-bound heme was included in the DG calculations through the addition of an artificial amino acid to the residue library used by the program DIANA.[13] The distance between the iron and the cyanide carbon was set to 1.85 Å, the Fe–CN

Figure 6.10 Number of (A) interresidual and (B) intraresidual NOEs per residue identified in the NOESY spectra of *S. c.* CN-Fe(III)Ala80cyt *c*. The total height of each column represents the total number of experimental NOEs. In (A), the black portions of the bars correspond to class 2 NOE constraints, the dark-shaded bars to class 3 NOE constraints, the light-shaded bars to class 4 and 5 NOE constraints, and the open bars to irrelevant constraints. In (B), open and filled bars indicate intraresidue NOE constraints which are irrelevant and meaningful, respectively. Residue 18 includes the heme as well as His18.


A
NOE per Residue
Residue
140
120
30
20
10
0
-5
10
20
30
40
50
60
70
80
90
100

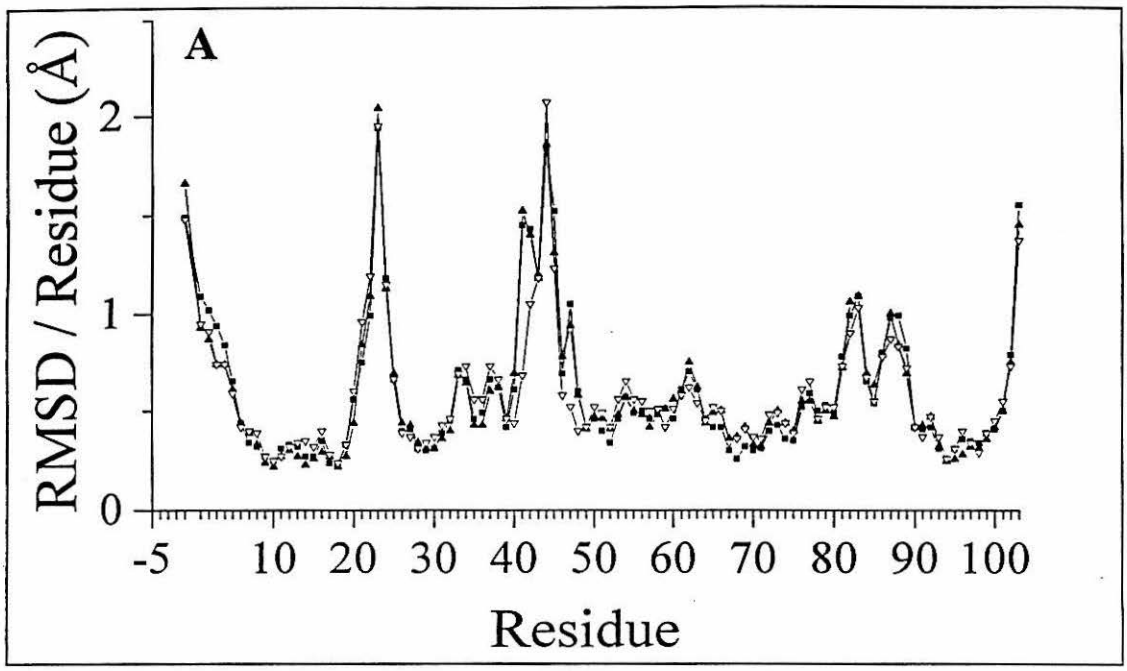

B
NOE per Residue
Residue
35
30
25
20
15
10
5
0
-5
10
20
30
40
50
60
70
80
90
100

distance measured in the X-ray structure of the cyanide adduct of cytochrome *c* peroxidase.[35] The artificial residue, denoted HES, consists of the heme-cyanide unit bound to a histidine residue, whose Hε2 is connected to the heme skeleton through links with the four pyrrole nitrogens (upper distance limit 2.9 Å). As the DIANA library does not allow the introduction of metal ions, the heme iron was defined as a dummy atom, and its hexacoordinate geometry set by imposing the appropriate upper and lower distance limits between the dummy atom and the ligand atoms by "special covalent bonds" (links). Starting from the four pyrrole nitrogens, the heme skeleton was built by defining normal covalent bonds between the atoms. "Special covalent bonds" were used to define the links between the heme and Cys14 and Cys17: 2.1 and 1.9 Å were used as upper and lower distance limits between each cysteine sulfur and the α-carbon of the corresponding heme thioether substituent. This yields a heme skeleton with the same distortion at that observed in the X-ray structure of *S. c.* Fe(III)Thr102cyt *c*.[30,36] The heme substituents have been allowed to assume all possible conformations by defining the appropriate bonds as rotatable. Analogously, the His18 ring was allowed to change its orientation with respect to the heme plane axes. Additional links between both carbon and nitrogen atoms of cyanide to the Nε2 of the His and the heme pyrrole nitrogens were introduced to maintain the cyanide ligand in position.

Distance Geometry Calculations

The DG family obtained using distance constraints derived from dipole-dipole connectivities consists of the 17 structures with the lowest target function (< 0.3 $Å^2$) and with residual violation of the distance constraints not exceeding 0.2 Å. The distribution of RMSD per residue among structures in the DG family (DG1) for both backbone and heavy atoms is shown in Figure 6.11. The average RMSD values within the DG1 family are 0.74 ± 0.10 and 1.35 ± 0.13 Å for backbone and all heavy atoms, respectively, (Table 6.5).

Figure 6.11 RMSD values per residue for the 17 accepted structures of the CN-Fe(III)Ala80cyt *c* for DG1 (filled squares), DG2 (filled triangles), and DG3 (inverted triangles). RMSD values for the backbone and heavy atoms are given in (A) and (B), respectively.


A
RMSD / Residue (Å)
0
1
2
-5
10
20
30
40
50
60
70
80
90
100
Residue

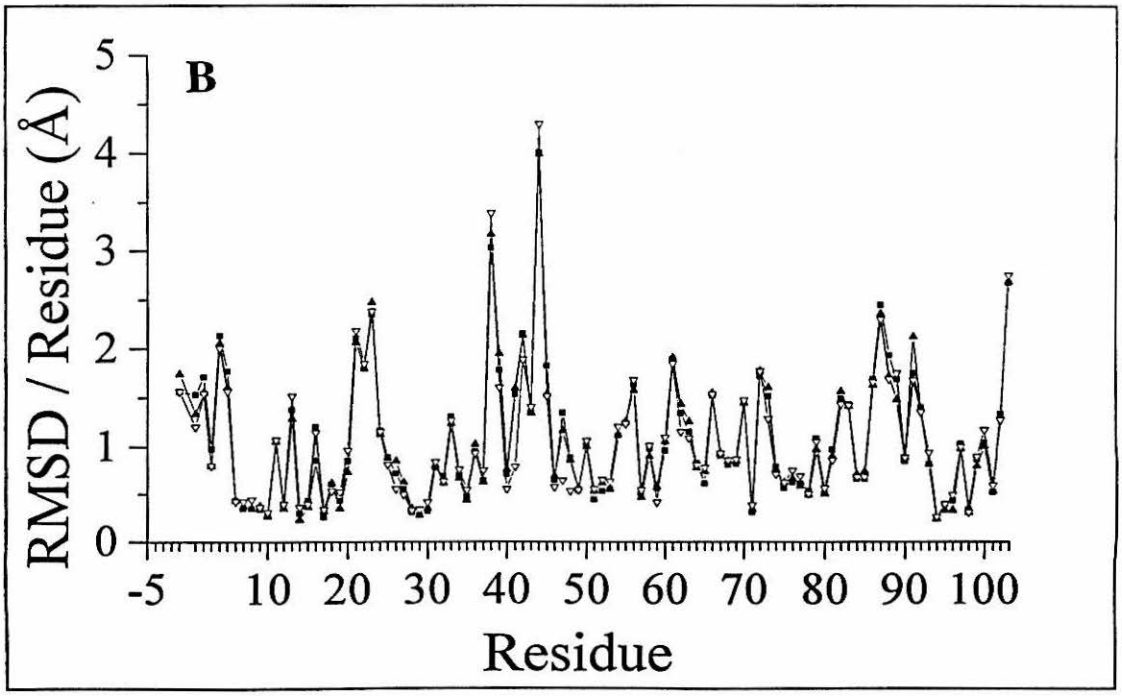

B
RMSD / Residue (Å)
0
1
2
3
4
5
-5
10
20
30
40
50
60
70
80
90
100
Residue

Table 6.5

Average RMSD values for the DG and REM families

	DG1[a]	DG2[b]	DG3[c]	REM
RMSD (residues 1–102, Å)				
Backbone	0.74 ± 0.10	0.71 ± 0.10	0.70 ± 0.11	0.68 ± 0.11
All heavy atoms	1.35 ± 0.13	1.34 ± 0.12	1.32 ± 0.13	1.32 ± 0.14
Target function ($Å^2$)	0.24	0.39	0.33	
total distance violations (Å)			3.9	4.2
total energy (kJ mol^{-1})				-6712.6
deviation from ideal bond distances (Å)				0.008
deviation from ideal bond angles (degrees)				1.60

[a]The 17 individual structures obtained from initial DG calculations.

[b]The 17 individual structures obtained from DG calculations excluding the constraints of the links between the heme and Cys14 and Cys17.

[c]The 17 individual structures obtained from DG calculations including hydrogen-bond constraints.

Two other DG calculations were then performed. The input data of the first (which generates family DG2) differ from that described above only in the lack of links between the heme and the sulfur atoms of Cys14 and Cys17. The removal of these heme-peptide linkages did not introduce any noticeable structural uncertainty within the set (see Table 6.5 and Figure 6.11) or any significant differences with respect to the DG1 family (the RMSD between the DG1 and DG2 average structures being ~ 0.2 Å the backbone and ~ 0.3 Å for all heavy atoms). This is a clear indication that the experimental NOESY constraints are sufficient to produce the correct protein folding and that the presence of links between Cys14 and Cys17 and the heme during the DG calculations does not bias the protein structure. A minor difference between families DG1 and DG2 is noted in the orientation of the Cys17 side chain and the heme 4-thioether β-CH_3.

A third DG calculation which included hydrogen-bond constraints was performed to produce family DG3. This approach is commonly used in cases when information on labile proton exchange rates is available to help define the solution structure.[33,37] The final output of the first DG calculation, which provided family DG1, was checked for possible hydrogen bonds involving both main-chain and side-chain atoms. The geometrical criteria used to indicate the presence of a hydrogen bond requires the distance between the hydrogen atom and the acceptor to be less than 2.4 Å and the angle formed by the donor atom, the H, and the acceptor atom to be $180 \pm 35°$. For the main-chain atoms, these geometrical criteria were found to be satisfied in a set of 39 calculated structures for the following 20 NH/CO pairs: Gly6/Lys4, Phe10/Gly6, Lys11/Ala7, Thr12/Leu9, Leu15/Phe10, Cys17/Cys14, His18/Cys14, Lys54/Asp50, Tyr67/Asn63, Leu68/Met64, Tyr74/Asn70, Ile75/Pro71, Leu94/Asp90, Ile95/Arg91, Thr96/Asn92, Tyr97/Asp93, Leu98/Leu94, Lys99/Ile95, Lys100/Thr96, Ala101/Thr97, and Ser102/Leu98. Hydrogen bonds are also present for the pairs His18 Hδ1/Pro30 CO, Val57 NH/Ser40 Oγ, Gly1 NH/Thr96 Oγ, and Ala43 NH/Tyr48 Oγ. Among these, the presence of hydrogen bonds involving the amide protons of residues 1, 10, 11, 12, 15, 43, 68, 75, 94, 97, 98, 99, 101, and 102 is supported by the experimental observation that the corresponding resonances are still present in the spectra recorded four weeks after the lyophilized sample was dissolved in D_2O solution. Slow exchange was experimentally observed also for the NH proton of the Trp59 indole ring. This proton falls within hydrogen-bond distance of the O(2) atom of heme 7-propionate. This O(2) atom falls within hydrogen-bonding distance of two other exchangeable protons, *i.e.*, Gly 41 NH and Asn 52 Hδ, but experimental evidence for slow exchange was found only for the amide proton of Gly41. All of these hydrogen bonds involving slowly exchanging protons were incorporated as constraints setting the distances between the H and the acceptor atom less than 2.4 Å and the distance between the donor and acceptor atoms in the range 2.7–3.4 Å (*i.e.*, to maintain an acceptable angle). For the Hε1 of Trp59 and the

O2 atom of heme 7-propionate, the donor-acceptor distance was allowed to range between 2.7 and 3.7 Å. The average RMSD values for the resulting family (DG3) are reported in Table 6.5, and the RMSD per residue is reported in Figure 6.11. Because a general improvement in the RMSD values is observed upon the introduction of the H-bond constraints, the DG3 family was used for structural refinements.

Restrained Energy Minimization and Restrained Molecular Dynamics

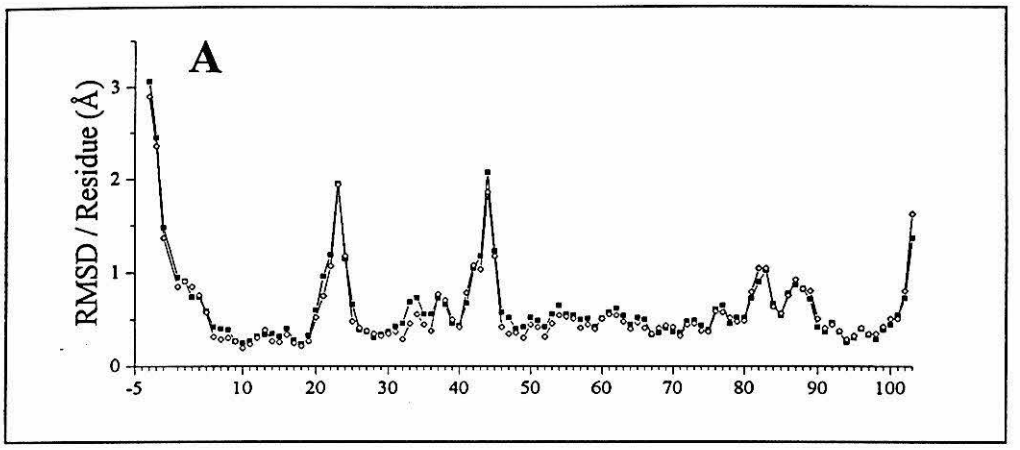
REM calculations were performed on the 17 structures of the DG3 family. The mean global RMSD values for residues 1−102 in the resulting REM family for the backbone and all heavy atoms are 0.68 ± 0.11 and 1.32 ± 0.14 Å, respectively (Table 6.5). In Figure 6.12, the distribution of the RMSD per residue (backbone and heavy atoms) is compared to that for the DG3 family. The stereo drawings of the DG3 and REM families (backbone, heme moiety, and His18 ligand) are shown in Figure 6.13.

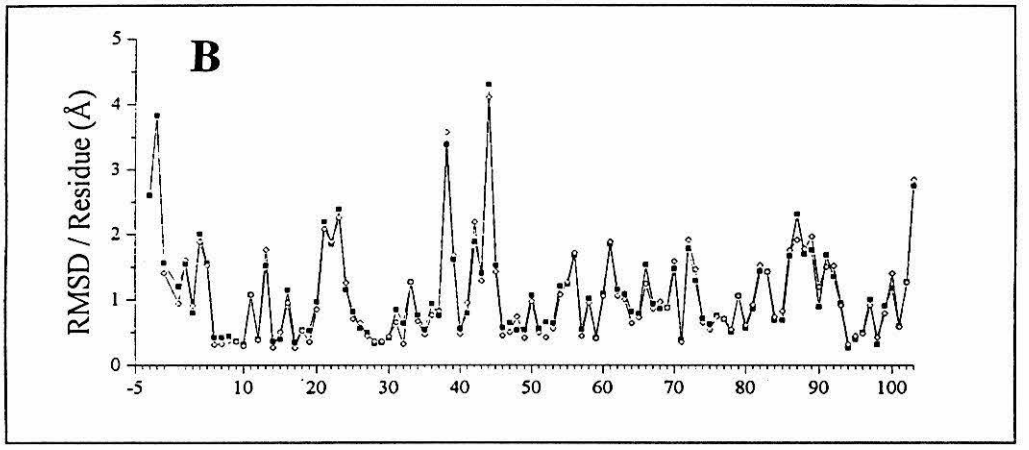
RMD calculations were performed on the REM family. The last 24 ps of the MD trajectory for each structure was used to generate an average structure which was then energy minimized to obtained the RMD family. The average RMSD values for the RMD family are 0.76 ± 0.12 and 1.45 ± 0.17 Å for the backbone and all heavy atoms, respectively, *i.e.*, somewhat higher than those of the REM family. It is possible that the force constant is not sufficient to provide a narrow family for the available number of constraints. This may be due to mobility and/or flexibility of a portion of protein, such as the N-terminus.

Evaluation of the Structure Quality

The RMSD values reported in Table 6.5 indicate that there is only a small improvement in the structure definition going from DG3 to REM. From inspection of Figure 6.13 it is evident that the backbone conformations are essentially the same in the two families.

Figure 6.12 RMSD per residue for the 17 accepted structures of *S. c.* CN-Fe(III)Ala80cyt *c* that constitute the DG3 (squares) and REM (diamonds) families. RMSD values for the backbone and all heavy atoms are reported in panels A and B, respectively. Panel C shows the average value of the distance violation per residue reported by SANDER after REM.


A
RMSD / Residue (Å)
0
1
2
3
-5
10
20
30
40
50
60
70
80
90
100


B
RMSD / Residue (Å)
0
1
2
3
4
5
-5
10
20
30
40
50
60
70
80
90
100

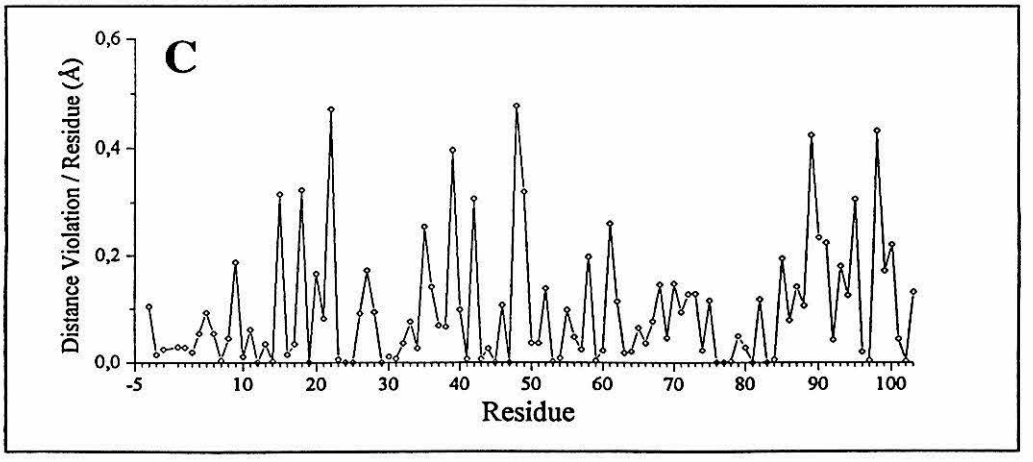
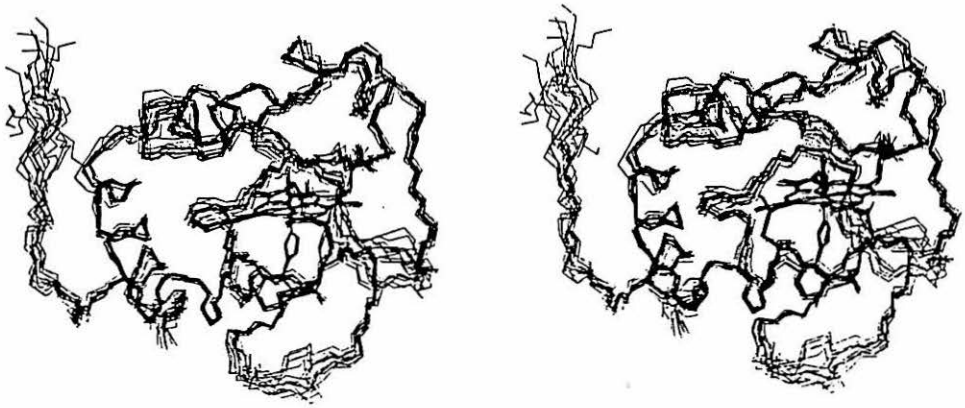
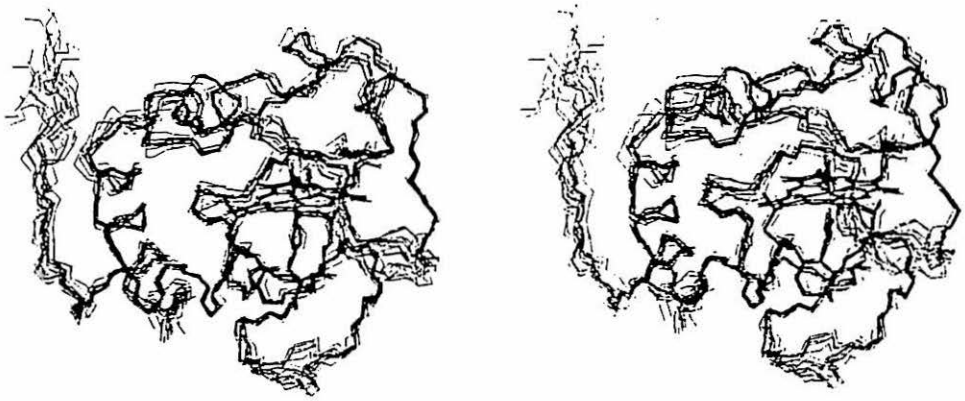

C
Distance Violation / Residue (Å)
0,0
0,2
0,4
0,6
-5
10
20
30
40
50
60
70
80
90
100
Residue

Figure 6.13 Stereo drawings of the 17 accepted structures of *S. c.* CN-Fe(III)Ala80cyt *c* that constitute the (A) DG3 and (B) REM families.

A



B



The presence of α-helices has been checked with the program PROCHECK.[27] In the two families DG3 and REM, residues 6–11, 50–55, 61–70, 71–75, and 89–101 have been found to form α-helices.

The N-terminal residues show relatively large RMSD values in all families. Substantial disorder for these residues has also been observed in the X-ray crystal structure of *S. c.* cyt *c* in both oxidation states[30,36] and may, at least in part, account for the lower structural resolution for DG3 compared to REM and for the increased disorder when passing to RMD. The residues involved in the first α-helix are characterized by low RMSD values in the DG3 and REM families (backbone RMSD of 0.40 Å for DG3 and 0.37 Å for REM).

Very low backbone RMSD values for the two families (0.32 and 0.27 Å in DG3 and REM) are found for residues 14–19, which include the heme-bound residues Cys14 and Cys17 and the iron-bound His18.

Residues 21–24 are characterized by larger RMSD values, with the largest being 1.95 and 1.94 Å for residues 23 in DG3 and REM, respectively. Indeed, the number of interresidue NOEs for these residues is low, averaging 10.2. RMD calculations drastically reduce the RMSD values (1.0 Å for residue 23).[38]

Residues 25–36 have RMSD values that are lower than average for the protein (backbone RMSD of 0.49 and 0.40 Å in DG3 and REM). Residues 37–47, on the contrary, exhibit the highest RMSD values in the structure. This protein region is indeed characterized by a low number of interresidue NOEs (9.6 on average). This is attributed to the structural properties of this segment of the protein; it does not have any defined secondary structural element and is projected toward the outside of the protein, forming an external loop. Therefore, few interresidue NOEs are expected.

Residues 48–80 are well defined in all three families, with RMSD values lower than average (0.49 and 0.48 Å for the backbone in DG3 and REM). As discussed above, three helices are identified (residues 50–55, 61–70, and 71–75) in this part of the protein.

Between residues 81 and 88, a slight increase in the spreading of the families (RMSD values for DG3 and REM being 0.80 and 0.83 Å) is observed. The relatively low number of interresidue NOEs in this part of the protein can be ascribed mainly to the lower number of assignments with respect to the other residues in the protein.

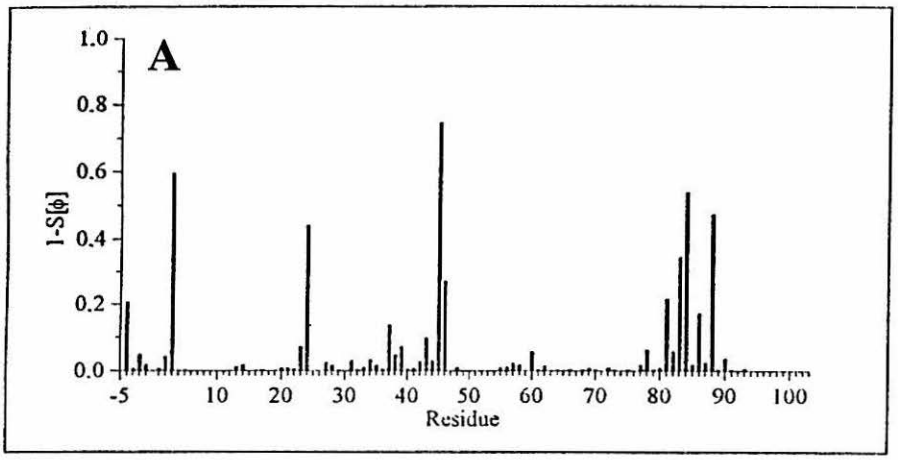
At the C-terminus, a long α-helix is present which is very well defined in all families. The calculated RMSD values for the backbone of residues 89–102 are 0.43 and 0.47 Å in DG3 and REM, respectively.

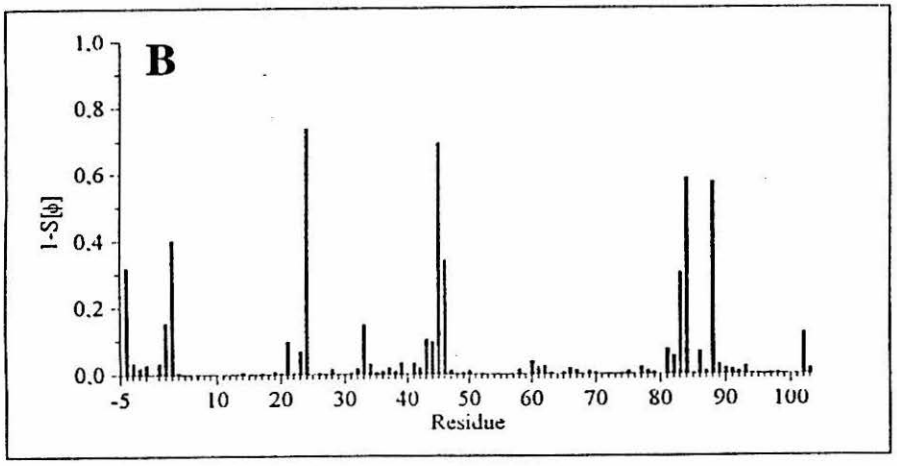
The trend of the RMSD values for the backbone atoms is observed also for all heavy atoms. The highest RMSD values are still found for residues 21–23, 38, 42, and 44. The 50–80 region, which is characterized by the lowest average RMSD values, shows larger fluctuations in these values with respect to the backbone atoms.

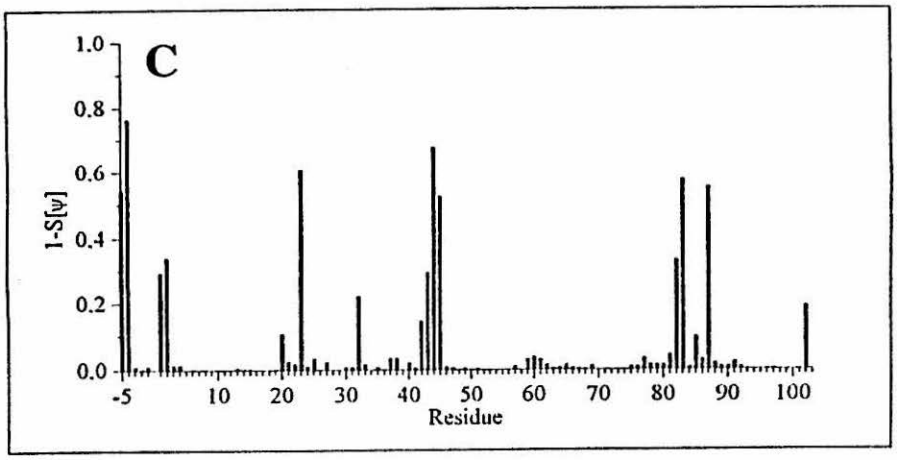
It is informative to compare the trend of the RMSD values per residue with the distance violation per residue found after the REM calculations (Figure 6.12C). Large penalty values do not correspond to violations in the DG calculations and are introduced by the coordinate variations which are allowed for the protein atoms during the energy minimization calculations. The total distance violation in the REM family is very low (4.2 Å), which corresponds to an average value per residue of 0.04 Å. All individual distance violations are lower than 0.28 Å. The largest penalty values are found for residues 22, 48, 89, and 98 (Figure 6.12C). The total penalty energy for these violations is 72.9 kJ mol^{-1}.

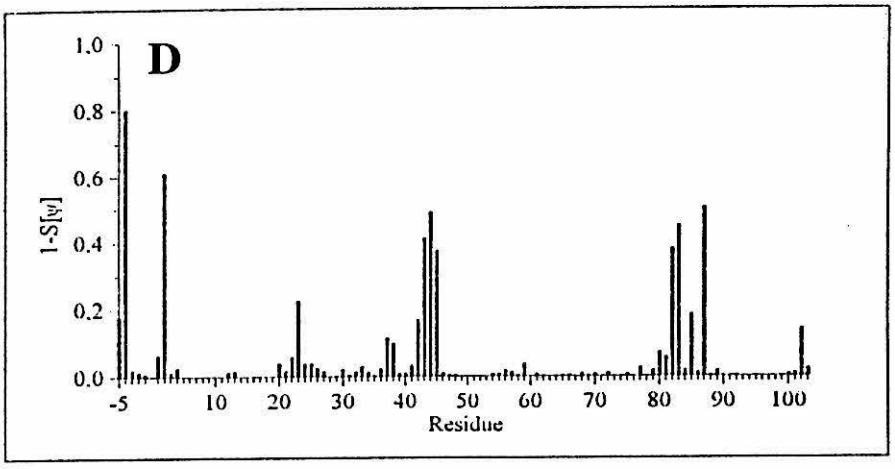
Another relevant parameter for the definition of the quality of the structure which can provide further insight in the analysis of the three families is given by the order parameter S. This parameter is 0 for randomly oriented structures and 1 for perfectly superimposed structures. The $1 - S$ quantity is conventionally used as measure of the disorder among structures within a family. Figure 6.14 reports the quantity $1 - S$ per residue as evaluated for the ϕ, ψ, and $\chi 1$ angles, *i.e.*, for the CO_{i-1}–N_i–C_i–CO_i (backbone), N_i–C_i–CO_i–N_{i+1} (backbone) and N_i–C_i–C_i–X_i (backbone to beginning

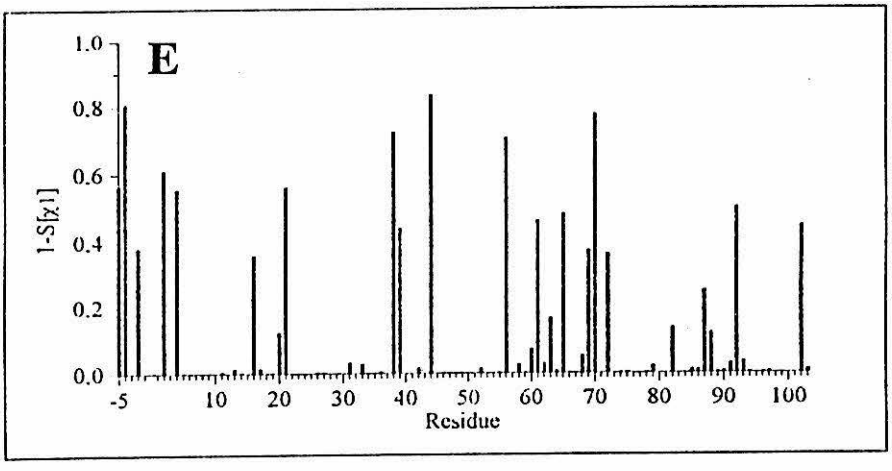
Figure 6.14 Histograms of the $1 - S$ parameter for the 17 structures of *S. c.* CN-Fe(III)Ala80cyt *c*. Panels (A), (C), and (E) indicate $1 - S$ for the DG3 family for the φ, ψ, and $\chi 1$ angles, respectively. Panels (B), (D), and (F) indicate the same for the REM family.

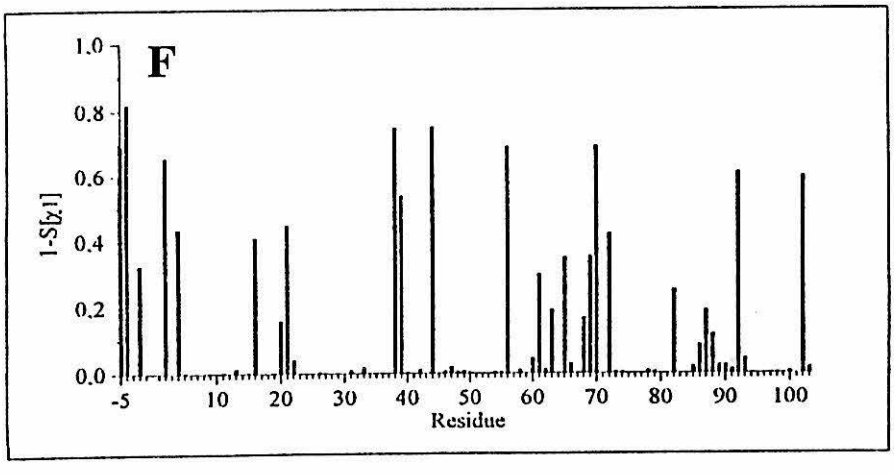

A
1-S[φ]
1.0
0.8
0.6
0.4
0.2
0.0
-5
10
20
30
40
50
60
70
80
90
100
Residue


B
1-S[φ]
1.0
0.8
0.6
0.4
0.2
0.0
-5
10
20
30
40
50
60
70
80
90
100
Residue


C
1-S[ψ]
1.0
0.8
0.6
0.4
0.2
0.0
-5
10
20
30
40
50
60
70
80
90
100
Residue


D
1-S[ψ]
1.0
0.8
0.6
0.4
0.2
0.0
-5
10
20
30
40
50
60
70
80
90
100
Residue


E
1-S[χ1]
1.0
0.8
0.6
0.4
0.2
0.0
-5
10
20
30
40
50
60
70
80
90
100
Residue


F
1-S[χ1]
1.0
0.8
0.6
0.4
0.2
0.0
-5
10
20
30
40
50
60
70
80
90
100
Residue

of the side chain) dihedral angles, respectively. For the ψ angle, a general improvement is observed going from DG3 to REM. This is particularly notable for residues 20–27. The $1 - S$ parameter for the φ and $\chi 1$ angles does not show meaningful changes between the two families, although there is a slight decrease in its absolute value.

The analysis of the Ramachandran plot, which correlates the ψ and φ dihedral angles, is another tool used to check the quality of the structures. In the REM average structure, 69 residues out of a total of 90 "meaningful" residues, *i.e.*, non-glycine, non-proline, and nonterminal residues (76.7% of all residues) fall in the most favored regions, with 18 residues (20.0%) in the additional allowed regions. Asn56 (1.1%) falls in a generously allowed region, while Lys5 and Glu88 (2.2%) are in disallowed regions. Given the resolution of the present structure (~ 2.5 Å according to a rough equivalence criterion between NMR and X-ray structures),[34] the percentage of residues in the most favored regions is in the range expected for a good quality structure.

Most of the side chains appear to have the same orientation in the DG3 and the REM families. Phe-3 is characterized by large RMSD values in both families. Phe-3 exhibits high disorder despite the large number (33) of interresidue NOEs observed for its ring protons. However, with the calibration used, most of them correspond to a very long upper distance limit, thus giving poor overall constraints for this residue. This suggests that the N-terminal portion of the protein is characterized by a larger mobility than the rest of the protein. As discussed above, this part of the protein chain also is characterized by large thermal factors in the X-ray structure.

Most of the charged side chains (Lys, Arg, Glu, Asp) are directed toward the protein exterior. They maintain the same orientation in DG3 and REM. An electrostatic interaction between Asp50 and Lys54 is observed in REM.

Particular attention should be devoted to the analysis of the definition of the structure of the heme, the heme-bound cysteines, the iron-bound histidine, and the residues in the "distal cavity" where the mutated Ala80 residue is present. Analysis of

the RMSD values (Figure 6.12A, B), the $1 - S$ parameters (Figure 6.14), and the distance violations (Figure 6.12C) for these residues indicates that this part of the protein is better defined than average. Of particular interest in this area of the protein is the Tyr67 residue, as it has been proposed to hydrogen-bond with heme-bound ligands. 1D NMR data, in particular the short T_1 value, large isotropic shift, and relatively slow exchange with solvent, suggested that this proton hydrogen-bonds with the heme-bound cyanide (see Chapter 5). In all of the calculated structures, this proton falls within hydrogen-bonding distance of the nitrogen atom of cyanide, although the angle requirement is not always fulfilled. The structural rearrangement of the Tyr67 ring in the REM family is shown in Figure 6.15. The hydroxyl proton, which is located 2.2 Å from the N atom of the cyanide ligand, has the same orientation for all members of the family. No differences are observed between the REM and RMD families for this residue.

Comparison with the Solution Structure of Horse Heart Cyt c

In Table 6.6, the average RMSD values between each component of the DG3 family with the mean DG3 structure <DG3>, are reported. The values of 0.47 ± 0.08 Å for the backbone and 0.91 ± 0.09 Å for all heavy atoms indicate that the structure quality is at the same level as that of the solution structure of horse heart Fe(II)cyt *c* (RMSD values about the average structure of 0.47 ± 0.09 and 0.91 ± 0.07 Å for the backbone and all heavy atoms, respectively).[33] The presence of a paramagnetic center, therefore, does not affect the overall resolution of the structure. Residues 23 and 24 are also reported to have the largest RMSD values in the solution structure of the horse heart protein.

Comparison with the X-ray Structure of Wild-Type Cyt c

S. c. CN-Ala80cyt *c* has the same overall fold as the wild-type protein. In Figure 6.16, the backbone trace of the average <REM> structure of the mutant is compared with the X-ray crystal structure of Fe(III)Thr102cyt *c*. The RMSD values between the X-ray

Figure 6.15 Stereo drawing of the active site of *S. c.* CN-Fe(III)Ala80cyt *c* given by the 17 structures of the REM family. The Tyr67 hydroxyl proton is also displayed to show the hydrogen bond with the nitrogen atom of cyanide. The bonds between the heme iron and the axial ligand are not displayed.

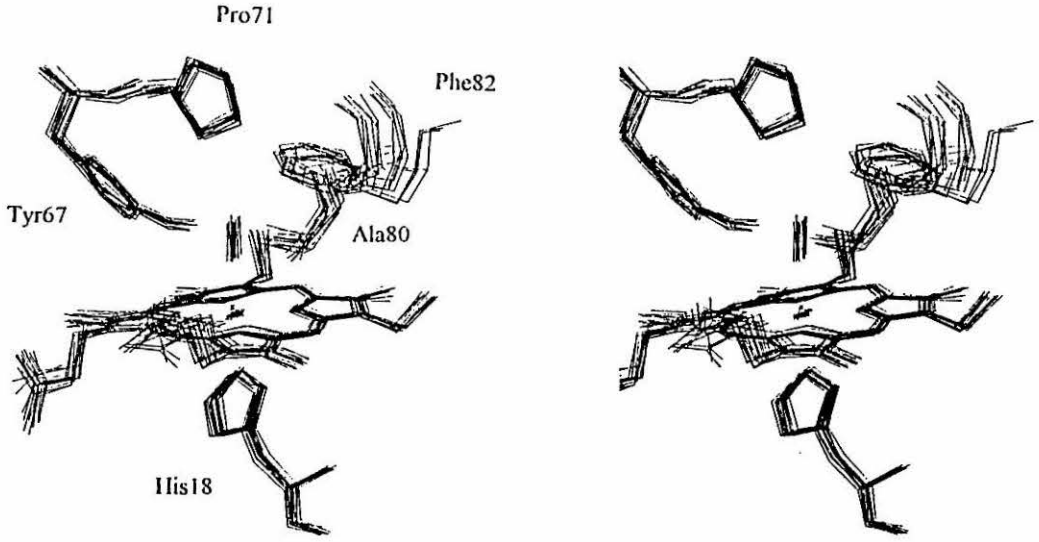

Pro71
Phe82
Tyr67
Ala80
His18

Table 6.6

Statistical analysis of the 17 calculated structures obtained by DG (DG3) and REM calculations[a]

	RMSD (Å)	
	backbone	all heavy atoms
DG3 vs. <DG3>	0.47 ± 0.08	0.91 ± 0.09
REM vs. <REM>	0.51 ± 0.07	0.94 ± 0.08
<DG3> vs. <REM>	0.48	0.55

[a]DG3 and REM refer to the structure families. <DG3> and <REM> are the mean structures obtained by averaging the coordinates of the respective families. RMSD values are for residues 1–102.

structure of *S. c.* cyt *c* in both oxidation states[30,36] and the average <DG3> and <REM> families of *S. c.* CN-Ala80 is reported in Table 6.7. The distribution of the RMSD for the backbone and all heavy atoms for the average <DG3> and <REM> structures with respect to the X-ray structure of oxidized *S. c.* cyt *c* is reported in Figure 6.17. The residues that deviate the most from the X-ray structure correspond to those with the largest RMSD values within the families (Figures 6.12, 6.13).

The side chains of the aromatic residues have similar orientations in the X-ray structure of WT and the solution structure of Ala80. The ring of Phe10 (which is very well defined in the solution structure with an RMSD value of 0.29 for the heavy atoms) has the same orientation but experiences a small translation. Phe82 is located above the heme ring in the "distal" pocket, on the solvent-exposed heme edge. It occupies an area between the heme and the solvent which is ~ 4 Å wide in the X-ray structure. In the solution structure, the ring of Phe82 has an average position closer to the heme but the movement is within the resolution of the residue. Also, the side chains of charged residues which have been proposed to be involved in interactions with physiological partners[39] have very similar orientations between the two structures. Significantly, very

Figure 6.16 Backbone atoms of (A) the mean REM structure of *S. c.* Fe(III)Ala80cyt *c* compared to (B) the X-ray crystal structure of *S. c.* Fe(III)Ser102cyt *c*. The heme and axial ligands are also shown in each structure.

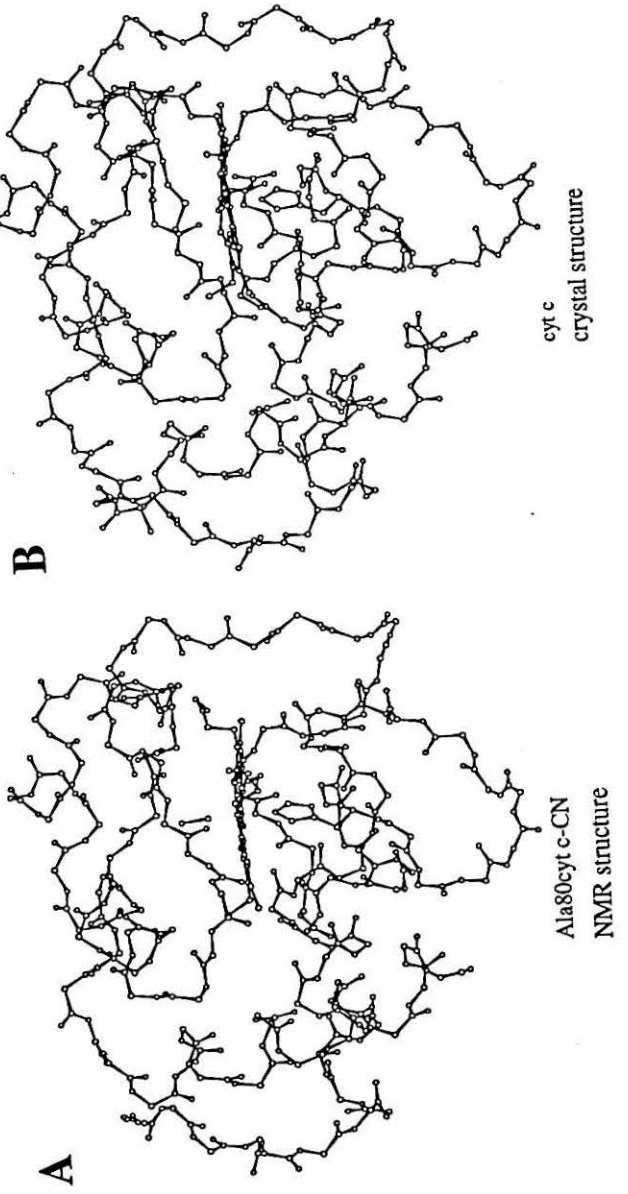
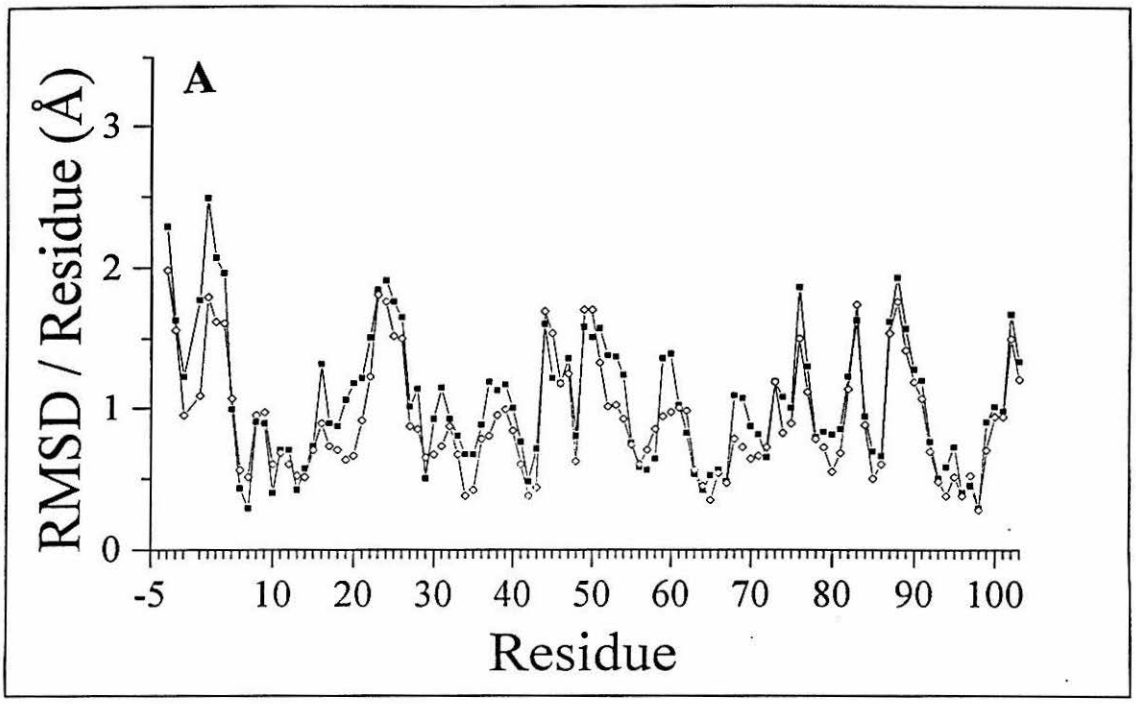

A
Ala80cyt c-CN
NMR structure
B
cyt c
crystal structure

Figure 6.17 RMSD values per residue for the average <DG3> (closed squares) and <REM> (open diamonds) structures with respect to the X-ray structure of *S. c.* Fe(III)Ser102cyt *c*. RMSD values for the backbone and all heavy atoms are given in (A) and (B), respectively. In the case of residue 39, the comparison with the X-ray structure is not meaningful because in the Ala80 variant has Gln in this position rather than His.


A
RMSD / Residue (Å)
0
1
2
3
-5
10
20
30
40
50
60
70
80
90
100
Residue

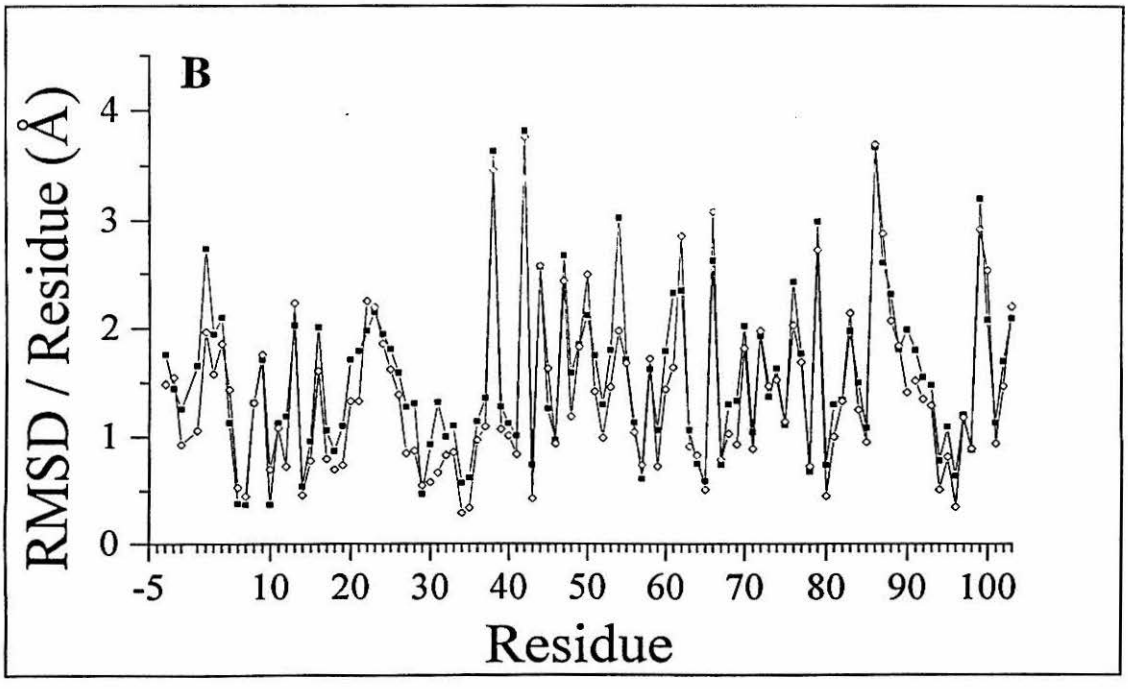

B
RMSD / Residue (Å)
0
1
2
3
4
-5
10
20
30
40
50
60
70
80
90
100
Residue

Table 6.7

RMSD values for the X-ray crystal structure of S. c. WTcyt c with respect to the <DG3> and <REM > structures of S. c. CN-Fe(III)Ala80cyt c

	RMSD (Å)[a]	
	<DG3>	<REM>
Fe(II)WTcyt *c*		
Backbone	1.10	0.95
All heavy atoms	1.64	1.50
Fe(III)WTcyt *c*		
Backbone	1.13	0.99
All heavy atoms	1.71	1.60

[a]RMSD for residues 1–102.

small RMSD values are observed for residue 18 (*i.e.*, the HES amino acid constituted by His18 and the heme), as well as for the two heme-bound cysteines (residues 14 and 17). The proximal His orientation observed in the solution structure is essentially unchanged from that observed in the X-ray structure of wild-type, consistent with the analysis of the heme methyl shifts (Chapter 5). Among the residues present in the "distal site," the backbone and heavy atoms of Pro71 experience very small RMSD values, indicating that they are in the same positions as observed in the X-ray structure. Therefore, although the shift values of the proton resonances of this residue are very different from those previously reported for *S. c.* Fe(III)Thr102cyt *c*,[31] they cannot be related to meaningful structural differences. Rather, they are due to different pseudocontact contributions to the hyperfine shifts arising from a different orientation of the magnetic axes.

For Tyr67, the RMSD value between the X-ray and the <REM> structure is also very small (0.47 and 0.79 Å for the backbone and all heavy atoms, respectively). This means that the location and orientation of this residue is similar to that observed in the X-ray structure of *S. c.* Thr102cyt *c*. However, based on ^{1}H NMR data reported for wild-

type yeast and horse cyt *c*,[31-33] the dynamics of this residue appear to be different. Tyr67 in wild-type forms a hydrogen bond with the sulfur atom of Met80, while in CN-Ala80 it forms a hydrogen bond with the cyanide moiety. The flipping rate of the ring is reduced in the mutant, changing from intermediate to slow on the NMR time scale compared to wild-type.

The substitution of the long side chain of Met80 with the short one of an alanine produces a large cavity on the "distal" side of the heme. The backbone of segment 79–81, which includes the modified Ala80 residue, is essentially the same in the two structures. As a result, a large distal cavity is created with no heavy atoms belonging to protein residues within ~ 6.5 Å of the iron. In myoglobin, the distal cavity is smaller, the iron having no heavy atoms of protein residues within ~ 5 Å,[40] while this radius in peroxidases is 5.5–6 Å.[35,41]

Conclusions

NMR spectroscopy has been shown to be a feasible method for the determination of the structure of a 108-amino acid paramagnetic heme protein in solution. This solution structure has the same resolution as that of reduced (diamagnetic) h. h. Fe(II)cyt *c*.[33] The paramagnetism of CN-Ala80 had both positive and negative effects on the sequence-specific assignment and, consequently, on the quality of the solution structure. Some of the residues that are strongly affected by the paramagnetism, such as His18 and Pro71, are among the most well-defined in the solution structure. This can be attributed to the large chemical shift dispersion and unusual chemical shifts seen for protons on these residues. Although this made their assignment nontrivial, it facilitated the identification and integration of interresidue NOESY cross peaks because they occurred in less crowded regions of the spectrum. On the other hand, the area of the protein including residues 81–85 is less well-defined than average and has fewer NOEs per residue. This is

due to the lower number of assignments and, thus, fewer identified NOESY cross peaks for these residues. In this case, the rapid nuclear relaxation caused by the paramagnetic center made the observation of cross peaks more difficult. Therefore, it is difficult to predict whether the presence of a paramagnetic center will increase or decrease the quality of a structure; as seen here, for a paramagnetic protein with a favorable (relatively fast) electron relaxation rate, the most likely case is a combination of effects.

One issue the solution structure helps to address is what factors account for the remarkable stability of the dioxygen adduct of *S. c.* Ala80cyt *c*. Structural modeling of the similar horse heart Ala80cyt *c* was used to provide an indication of what kinds of distal pocket-dioxygen interactions are possible (Chapter 2). One of these proposed interactions is that between the dioxygen O(2), which has partial negative charge, and the positive edge of the Phe82 ring. The solution structure shows that the Phe82 δ-carbon is 3.3 Å from the cyanide nitrogen, which is very similar to the 3.2 Å distance between the O(2) atom of dioxygen and the ζ-carbon of Phe29 (position B10) in the X-ray structure of Leu29Phe sperm whale myoglobin, which has an unusually slow autoxidation rate.[42] The solution structure supports the possibility that this interaction occurs in *S. c.* Ala80cyt *c*. This type of interaction involving aromatic rings has now been proposed for a number of naturally occurring Mbs and Hbs. Asian elephant Mb has a Gln in place of the distal His. Although capable of hydrogen-bonding, it is not thought to be as effective as a His in this capacity.[43] However, a position B10 Phe has been identified in this Mb and NMR and X-ray data indicate that it interacts with the heme-dioxygen unit.[44] Hb I from *Lucina pectinata* is characterized by a Gln in position E7 and Phe residues in both the B10 and E11 positions. The X-ray crystal structure of the oxidized form of this Hb indicates that these Phe's may interact with heme-bound ligands.[45]

The major distal cavity interaction involved in stabilization of Ala80cyt *c* is proposed to be hydrogen bonding from the Tyr67 OH to the iron-bound dioxygen, in

analogy to the hydrogen-bonding interaction seen between the distal His and dioxygen in Mb. Experimental support for this interaction is now given by the solution structure and the slow exchange with solvent exhibited by the Tyr67 OH. Other residues, in particular Gln[43-46] and Arg,[47] have also been shown to participate in this type of interaction in various Mb species. Of particular interest is the report of hydrogen bonding between the heme-bound dioxygen O(2) and a position B10 tyrosine in *Ascaris* Hb.[46] This Hb has both an exceptionally high dioxygen binding constant and a low autoxidation rate.[48] This interaction therefore may account both for much of the strong dioxygen binding and the resistance of the dioxygen adduct of Ala80cyt *c* to autoxidation. The dioxygen affinity of sperm whale Mb has in fact been increased significantly by protein engineering by substituting Leu29 (B10) with Tyr.[49]

The presence of a distal pocket Tyr, however, does not necessarily lead to formation of a stable dioxygen adduct.[50] In Ala80cyt *c*, an important component of dioxygen stabilization is related to the protection of the iron site by the protein matrix. This is observed experimentally by the usually slow rate of CO binding to h. h. Fe(II)Ala80 as well as the very low quantum yields of photoejection of both CO and O_2 from h. h. and *S. c.* Fe(II)Ala80. The protection of the heme from solvent slows autoxidation by preventing release of superoxide (or dioxygen by an alternative mechanism, see Chapter 2) and subsequent binding of water/hydroxide as required by the autoxidation mechanism as observed in Mb.[51] The similarity of the overall structure of Ala80cyt *c* to that of the native protein confirms that the cytochrome fold has structural and/or dynamic properties that hinder the movement of ligands between the heme and solution. As an electron transfer protein, it is important for cyt *c* to prevent exogenous ligand binding as this would have the effect of precluding facile electron transfer. The remarkable feature of Mbs and Hbs in nature is that they are able to form stable dioxygen adducts while still allowing efficient uptake and release of ligands.

References and Notes

(1) The work in this chapter was done in collaboration with Dr. Paola Turano, Pornthep Sompornpsiut, Prof. Ivano Bertini, and Prof. Lucia Banci at the University of Florence. The coordinates of the family of 17 structures constituting the REM family is available in the Brookhaven Protein Data Bank (identifier 1FHB).

(2) The first X-ray crystal structures of proteins were of myoglobin (Kendrew, J. *Science* **1963,** *139,* 1259-1266) and of hemoglobin (Perutz, M. F. *Science* **1963,** *140,* 863-869).

(3) Williamson, M. P.; Havel, T. J.; Wüthrich, K. *J. Mol. Biol.* **1985,** *182,* 295-315.

(4) *Macromolecular Structures*; Hendrickson, W. A., Wüthrich, K., Eds. Current Biology: London, 1995.

(5) For recent reviews comparing protein structure determination by NMR and crystallography, see: Wüthrich, K. *Acta Cryst. D* **1995,** *51,* 249-270. Wagner, G.; Hyberts, S. G.; Havel, T. F. *Ann. Rev. Biophys. Biomol. Struct.* **1992,** *21,* 167-198.

(6) Cistola, D. P.; Hall, K. B. *J. Biomol. NMR.* **1995,** *5,* 415-419. Clore, G. M.; Bax, A.; Wingfield, P. T.; Gronenborn, A. M. *Biochemistry* **1990,** *29,* 5671-5675.

(7) Protein structure determination by three- and four-dimensional NMR spectroscopy is reviewed in: Oschkinat, H.; Muller, T.; Dieckmann, T. *Ang. Chem. Int. Ed. Eng.* **1994,** *33,* 277-293.

(8) Bertini, I.; Luchinat, C. *NMR of Paramagnetic Molecules in Biological Systems*; Benjamin/Cummings: Menlo Park, 1986.

(9) Bertini, I.; Turano, P.; Vila, A. J. *Chem. Rev.* **1993,** *93,* 2833-2932.

(10) Techniques used for NMR spectroscopy of paramagnetic proteins are reviewed in: Banci, L.; Piccioli, M.; Scozzafava, A. *Coord. Chem. Rev.* **1992,** *120,* 1-28. Bertini, I.; Capozzi, F.; Luchinat, C.; Turano, P. *J. Mag. Reson.* **1991,** *95,* 244-252.

(11) Banci, L.; Bertini, I.; Eltis, L. D.; Felli, I. C.; Kastrau, D. H. W.; Luchinat, C.; Piccioli, M.; Pierattelli, R.; Smith, M. *Eur. J. Biochem.* **1994,** *225,* 715-725.

(12) Reduced *Chromatium vinosum* high-potential iron-sulfur protein: Banci, L.; Bertini, I.; Dikiy, A.; Kastrau, D. H. W.; Luchinat, C.; Sompornpisut, P. *Biochemistry* **1995,** *34,* 206-219. Oxidized *Chromatium vinosum* high-potential iron-sulfur protein: Bertini, I.; Dikiy, A.; Kastrau, D. H. W.; Luchinat, C.; Sompornpisut, P. *Biochemistry* **1995,** *34,* 9851-9858. Oxidized *Clostridium pasteurianum* 2[4Fe-4S] ferredoxin: Bertini, I.; Donaire, A.; Feinberg, B. A.; Luchinat, C.; Piccioli, M.; Yuan, H. P. *Eur. J. Biochem.* **1995,** *232,* 192-205.

(13) Güntert, P.; Braun, W.; Wüthrich, K. *J. Mol. Biol.* **1991,** *217,* 517-530.

(14) Wüthrich, K. *Acc. Chem. Res.* **1989,** *22,* 36-44.

(15) Güntert, P.; Wüthrich, K. *J. Biomol. NMR* **1991,** *1,* 447-456.

(16) Eccles, C.; Güntert, P.; Billeter, M.; Wüthrich, K. *J. Biomol. NMR* **1991,** *1,* 111-130.

(17) Güntert, P.; Berndt, K. D.; Wüthrich, K. *J. Biomol. NMR* **1993,** *3,* 601-606.

(18) Pearlman, D. A.; Case, D. A.; Caldwell, G. C.; Siebel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. *AMBER 4.0*; University of California: San Francisco, 1991.

(19) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comp. Chem.* **1986,** *7,* 287-303.

(20) Banci, L.; Carloni, P.; Gori Savelini, G. *Biochemistry* **1994,** *33,* 12356-12366.

(21) Angelucci, L.; De Gioia, L.; Fantucci, P. *Gazz. Chim. Ital.* **1993,** *123,* 111-117. Dawson, J. H.; Andersson, L. A.; Sono, M. *J. Biol. Chem.* **1982,** *257,* 3606-3617. Case, D. A.; Karplus, M. *J. Mol. Biol.* **1979,** *132,* 343-368. Ruf, H. H.; Wende, P.; Ullrich, V. *J. Inorg. Biochem.* **1979,** *11,* 1890-204. Argos, P.; Mathews, F. S. *J. Biol. Chem.* **1975,** *250,* 747-751.

(22) Jones, L. H. *Inorganic Vibrational Spectroscopy*; Marcel-Dekker: New York, 1971. Hancock, R. D. In *Progress in Inorganic Chemistry*; Lippard, S. J., Ed.; John Wiley & Sons, Inc.: New York, 1989; Vol. 37, p 187. Nakamoto, K.; Takemoto, J.; Chow, T. L. *Appl. Spectrosc.* **1971,** *25,* 352-355. Nakamoto, K. *Infrared and Raman Spectra of Inorganic and Coordination Compounds*; John Wiley & Sons: New York, 1978.

(23) F. J. Seiler Research Laboratory, U. S. Air Force Academy, Colorado, Spring, CO.

(24) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. *J. Chem. Phys.* **1984,** *81,* 3684-3690.

(25) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977,** *34,* 1311-1327.

(26) Ross, W. S. *CARNAL*; Department of Pharmaceutical Chemistry: University of California, San Francisco, 1994.

(27) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallograph.* **1993,** *26,* 283-291.

(28) Sayle, R. Biomolecular Structure Department, Glaxo Rsearch and Development, Greenford, Middlesex. U.K., 1994.

(29) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.

(30) Louie, G. V.; Hutcheon, W. L. B.; Brayer, G. D. *J. Mol. Biol.* **1988,** *199,* 295-314. Louie, G. V.; Brayer, D. G. *J. Mol. Biol.* **1990,** *214,* 527-555.

(31) Gao, Y.; Boyd, J.; Williams, R. J. P.; Pielak, G. J. *Biochemistry* **1990,** *29,* 6994-7003.

(32) Moore, G. R.; Williams, R. H. P. *Eur. J. Biochem.* **1980,** *103,* 513-522. Moore, G. R.; Williams, R. H. P. *Eur. J. Biochem.* **1980,** *103,* 523-532. Wand, A. J.; Di Stefano, D. L.; Feng, Y.; Roder, H.; Englander, S. W. *Biochemistry* **1989,** *28,* 186-194. Wand, A. J.; Di Stefano, D. L.; Feng, Y.; Roder, H.; Englander, S. W. *Biochemistry* **1989,** *28,* 195-203.

(33) Qi, P. X.; Di Stefano, D. L.; Wand, A. J. *Biochemistry* **1994,** *33,* 6408-6417.

(34) Clore, G. M.; Gronenborn, A. M. *Protein Sci.* **1994,** *3,* 372-390.

(35) Edwards, S. L.; Poulos, T. L. *J. Biol. Chem.* **1990,** *265,* 2588-2595.

(36) Berghuis, A. M.; Brayer, G. D. *J. Mol. Biol.* **1992,** *223,* 959-976.

(37) Detlefsen, D. J.; Thanabal, V.; Pecoraro, V. L.; Wagner, G. *Biochemistry* **1991,** *30,* 9040-9046. Markus, M. A.; Nakayama, T.; Matsudaira, P.; Wagner, G. *Protein Sci.* **1994,** *3,* 70-81.

(38) The spreading of the structures in this region of the protein in the RMD family is drastically reduced with respect to the REM and DG families. The structures, therefore, even if not well defined with respect to NOEs, are able to find conformations at minimum energy that are similar to each other during MD calculations.

(39) Moore, G. R.; Pettigrew, G. W. *Cytochromes c; Evolutionary, Structural and Physicochemical Aspects*; Springer-Verlag: Berlin, 1990. Pettigrew, G. W.; Moore, G. R. *Cytochromes c; Biological Aspects*; Springer-Verlag: Berlin, 1987.

(40) Kuriyan, J.; Wilz, S.; Karplus, M.; Petsko, G. A. *J. Mol. Biol.* **1986,** *192,* 133-154.

(41) Edwards, S. L.; Raag, R.; Wariishi, H.; Gold, M. H.; Poulos, T. L. *Proc. Natl. Acad. Sci. U.S.A.* **1993,** *90,* 750-754. Smith, T. D.; Gaunt, R.; Ruzic, I. *Inorg. Chim. Acta* **1983,** *78,* 103-106. Poulos, T. L.; Kraut, J. *J. Biol. Chem.* **1980,** *255,* 8199-8205. Finzel, B. C.; Poulos, T. L.; Kraut, J. *J. Biol. Chem.* **1984,** *259,* 13027-13036.

(42) Carver, T. E.; Brantley, R. E., Jr.; Singleton, E. W.; Arduini, R. M.; Quillin, M. L.; Phillips, G. N., Jr.; Olson, J. S. *J. Biol. Chem.* **1992,** *267,* 14443-14450.

(43) Springer, B. A.; Sligar, S. G.; Olson, J. S.; Phillips, G. N. *Chem. Rev.* **1994,** *94,* 699-714.

(44) Vyas, K.; Rajarathnam, K.; Yu, L. P.; Emerson, S. D.; La Mar, G. N.; Krishnamoorthi, R.; Mizukami, H. *J. Biol. Chem.* **1993,** *268,* 14826-14835.

Bisig, D. A.; Di Iorio, E. E.; Diederichs, K.; Winterhalter, K. H.; Piontek, K. *J. Biol. Chem.* **1995,** *270,* 20754-20762.

(45) Rizzi, M.; Wittenberg, J. B.; Coda, A.; Fasano, M.; Ascenzi, P.; Bolognesi, M. *J. Mol. Biol.* **1994,** *244,* 86-99.

(46) De Baere, I.; Perutz, M. F.; Kiger, L.; Marden, M. C.; Poyart, C. *Proc. Natl. Acad. Sci. U.S.A.* **1994,** *91,* 1594-1597. Kloek, A.; Yang, J.; Mathews, F. S.; Frieden, C.; Goldberg, D. E. *J. Biol. Chem.* **1994,** *269,* 2377-2379.

(47) Qin, J.; La Mar, G. N.; Ascoli, F.; Bolognesi, M.; Brunori, M. *J. Mol. Biol.* **1992,** *224,* 891-897. Yamamoto, Y.; Iwafune, K.; Chûjo, R.; Inoue, Y.; Imai, K.; Suzuki, T. *J. Mol. Biol.* **1992,** *228,* 343-346.

(48) Gibson, Q. H.; Smith, M. *Proc. R. Soc. Lond. B* **1965,** *163,* 206-214. Okazaki, T.; Wittenberg, B. A.; Briehl, R. W.; Wittenberg, J. B. *Biochim. Biophys. Acta* **1967,** *111,* 485-495.

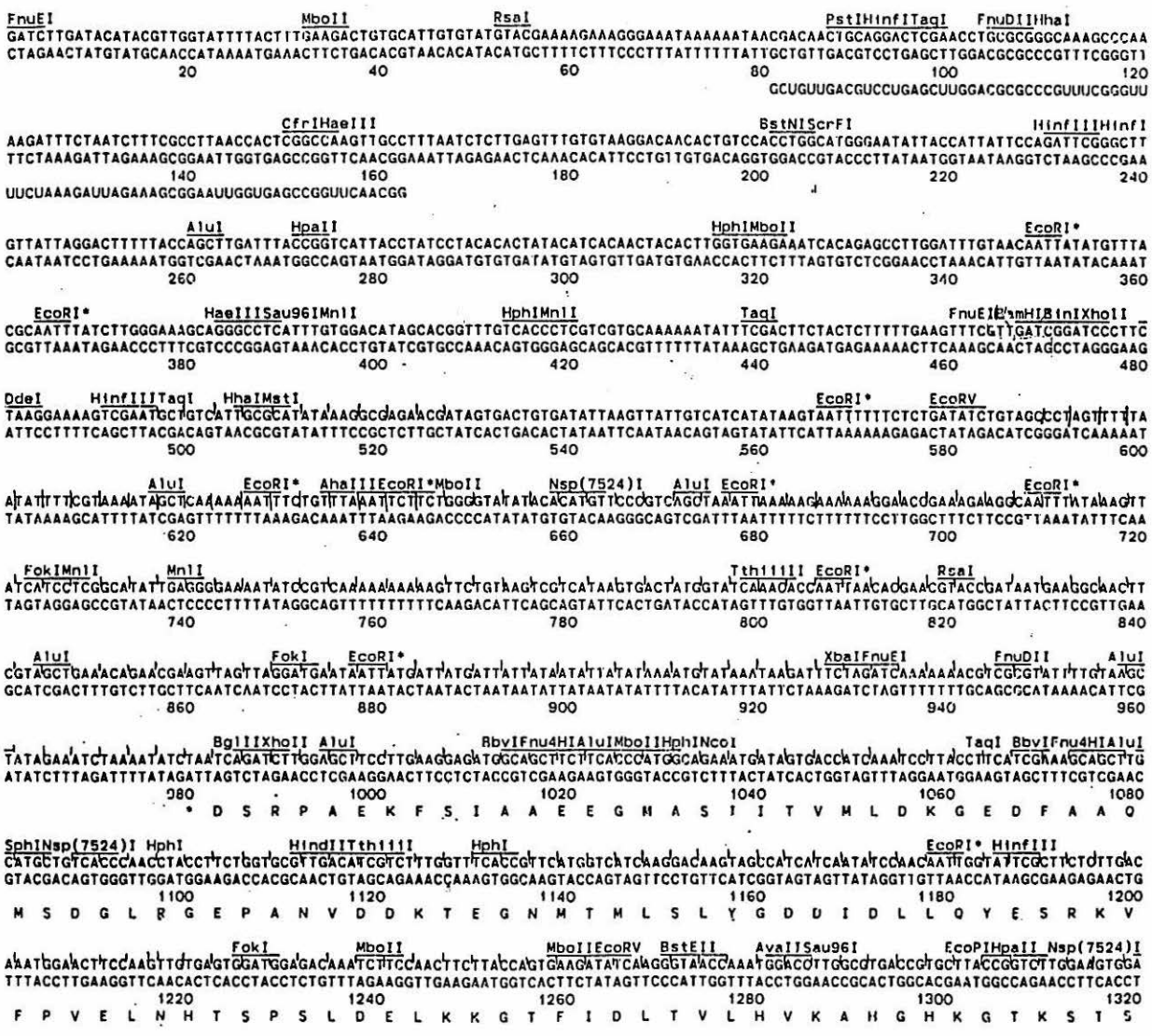
(49) Gibson, Q. H.; Regan, R.; Olson, J. S.; Carver, T. E.; Dixon, B.; Pohajdak, B.; Sharma, P. K.; Vinogradov, S. N. *J. Biol. Chem.* **1993,** *268,* 16993-16998. Travaglini Allocatelli, C.; Cutruzzolà, F.; Brancaccio, A.; Vallone, B.; Brunori, M. *FEBS Lett.* **1994,** *352,* 63-66.

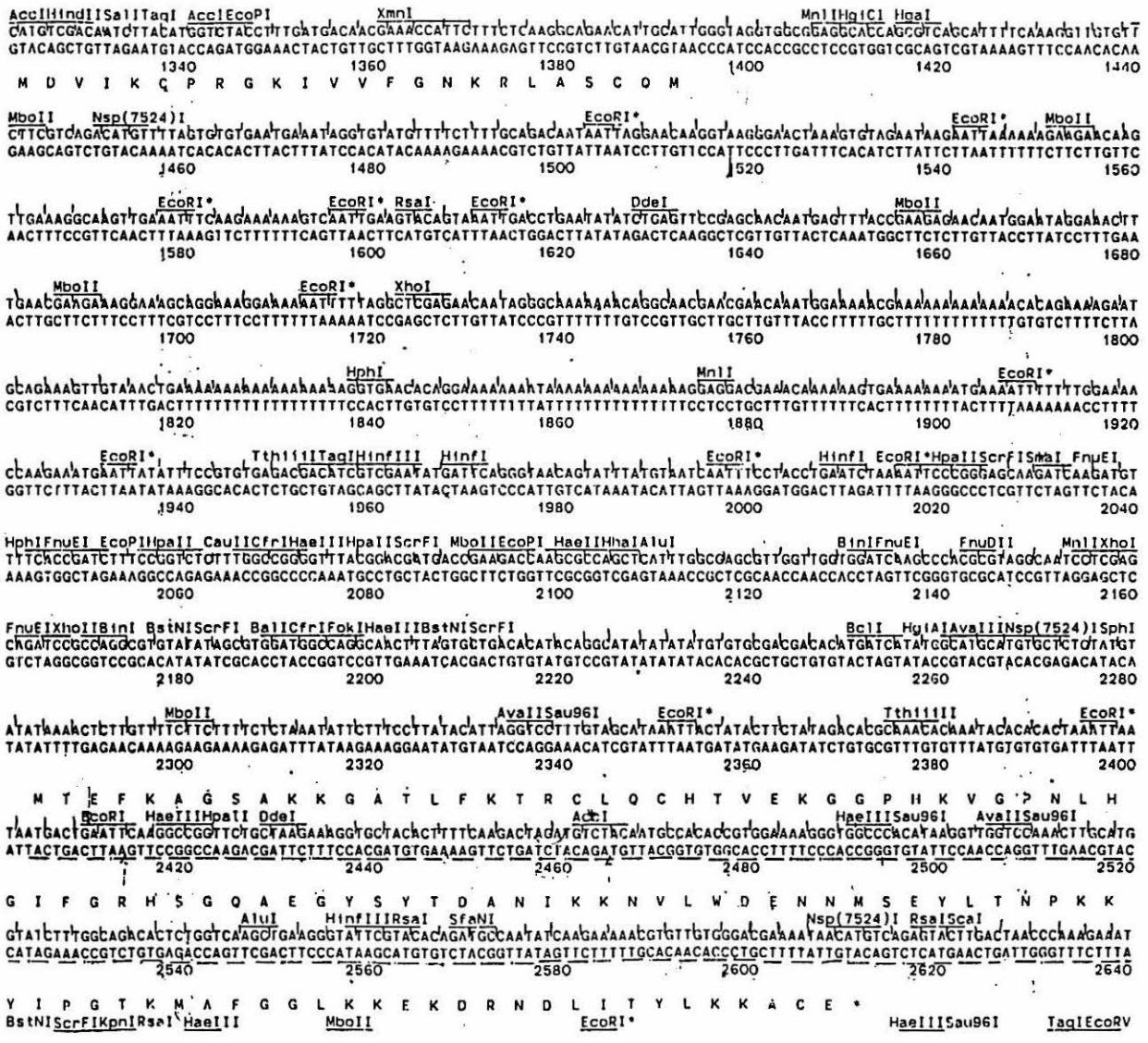
(50) His64Tyr Mb is characterized by a very rapid autoxidation rate: Springer, B. A.; Egeberg, K. D.; Sligar, S. G.; Rohlfs, R. J.; Mathews, A. J.; Olson, J. S. *J. Biol. Chem.* **1989,** *264,* 3057-3060.

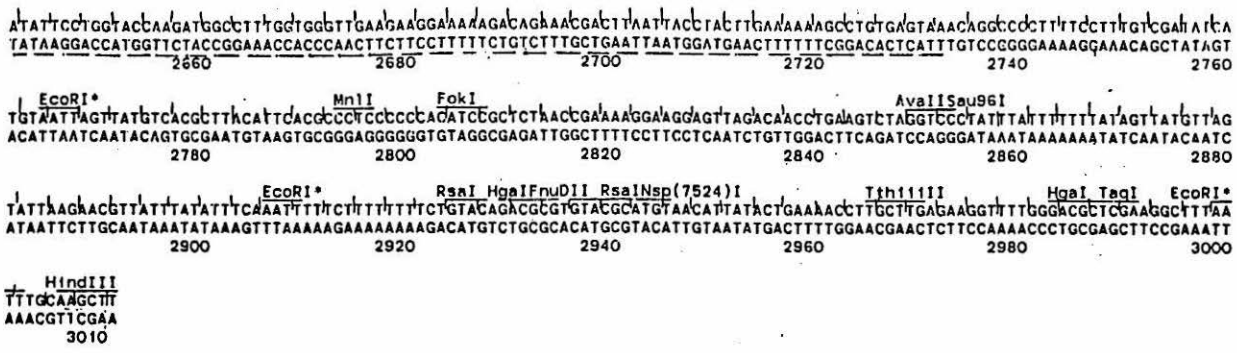
(51) Brantley, R. E.; Smerdon, S. J.; Wilkinson, A. J.; Singleton, E. W.; Olson, J. S. *J. Biol. Chem.* **1993,** *168,* 6995-7010.

Appendix A

Sequence of 3-kb DNA Fragment Containing the *CYC1* Gene


FnuEI MboII RsaI PstIHinfITaqI FnuDIIHhaI
GATCTTGATACATACGTTGGTATTTTACTTTGAAGACTGTGCATTGTGTATGTACGAAAAGAAAGGGAAATAAAAAATAACGACAACTGCAGGACTCGAACCTGCGCGGGCAAAGCCCAA
CTAGAACTATGTATGCAACCATAAAATGAAACTTCTGACACGTAACACATACATGCTTTTCTTTCCCTTTATTTTTTATTGCTGTTGACGTCCTGAGCTTGGACGCGCCCGTTTCGGGTT
20 40 60 80 100 120
GCUGUUGACGUCCUGAGCUUGGACGCGCCCGUUUCGGGUU
CfrIHaeIII BstNIScrFI HinfIIIHinfI
AAGATTTCTAATCTTTCGCCTTAACCACTCGGCCAAGTTGCCTTTAATCTCTTGAGTTTGTGTAAGGACAACACTGTCCACCTGGCATGGGAATATTACCATTATTCCAGATTCGGGCTT
TTCTAAAGATTAGAAAGCGGAATTGGTGAGCCGGTTCAACGGAAATTAGAGAACTCAAACACATTCCTGTTGTGACAGGTGGACCGTACCCTTATAATGGTAATAAGGTCTAAGCCCGAA
140 160 180 200 220 240
UUCUAAAGAUUAGAAAGCGGAAUUGGUGAGCCGGUUCAACGG
AluI HpaII HphIMboII EcoRI*
GTTATTAGGACTTTTTACCAGCTTGATTTACCGGTCATTACCTATCCTACACACTATACATCACAACTACACTTGGTGAAGAAATCACAGAGCCTTGGATTTGTAACAATTATATGTTTA
CAATAATCCTGAAAAATGGTCGAACTAAATGGCCAGTAATGGATAGGATGTGTGATATGTAGTGTTGATGTGAACCACTTCTTTAGTGTCTCGGAACCTAAACATTGTTAATATACAAAT
260 280 300 320 340 360
EcoRI* HaeIIISau96IMnlI HphIMnlI TaqI FnuEIBamHIBinIXhoII
CGCAATTTATCTTGGGAAAGCAGGGCCTCATTTGTGGACATAGCACGGTTTGTCACCCTCGTCGTGCAAAAAATATTTCGACTTCTACTCTTTTTGAAGTTTCCTTGATCGGATCCCTTC
GCGTTAAATAGAACCCTTTCGTCCCGGAGTAAACACCTGTATCGTGCCAAACAGTGGGAGCAGCACGTTTTTTATAAAGCTGAAGATGAGAAAAACTTCAAAGCAACTAGCCTAGGGAAG
380 400 420 440 460 480
DdeI HinfIIITaqI HhaIMstI EcoRI* EcoRV
TAAGGAAAAGTCGAATGCTGTCATTGCGCATATAAAGGCGAGACGATAGTGACTGTGATATTAAGTTATTGTCATCATATAAGTAATTTTTTCTCTGATATCTGTAGCCCTAGTTTTTA
ATTCCTTTTCAGCTTACGACAGTAACGCGTATATTTCCGCTCTTGCTATCACTGACACTATAATTCAATAACAGTAGTATATTCATTAAAAAAGAGACTATAGACATCGGGATCAAAAAT
500 520 540 560 580 600
AluI EcoRI* AhaIIIEcoRI*MboII Nsp(7524)I AluI EcoRI* EcoRI*
ATATTTTCGTAAATATAGCTCAAAAAATTTCTGTTTAAATTCTTCTGGGGTATATACACATGTTCCGGTCAGGTAAATTAAATAAGAAAAAGGAACGGAATAGAAGGCAATTTTATAAGTT
TATAAAAGCATTTTATCGAGTTTTTTAAAGACAAATTTAAGAAGACCCCATATATGTGTACAAGGGCAGTCGATTTAATTTTCTTTTTTCCTTGGCTTTCTTCCGTTAAATATTTCAA
620 640 660 680 700 720
FokIMnlI MnlI Tth111I EcoRI* RsaI
ATCATCCTCGGCATATTGAGGGGAAAATATCCGTCAAAAAAAAAAGTTCTGTAAGTCGTCATAATGTGACTATGGTATCAAAACACCAATTAACAGGAACGTACCGATAATGAAGGCAACTT
TAGTAGGAGCCGTATAACTCCCCTTTTATAGGCAGTTTTTTTTTTCAAGACATTCAGCAGTATTCACTGATACCATAGTTTGTGGTTAATTGTGCTTGCATGGCTATTACTTCCGTTGAA
740 760 780 800 820 840
AluI FokI EcoRI* XbaIFnuEI FnuDII AluI
CGTAGCTGAAACAGAACGAAGTTAGTTAGGATGAATAATTATGATTATGATTATTATATATTATATTAAATATGTATAAATAAGATTTCTAGATCAAAAAACGTCGTGTATTTTGTAAGC
GCATCGACTTTGTCTTGCTTCAATCAATCCTACTTATTAATACTAATACTAATAATATTATAATATATTTTACATATTTATTCTAAAGATCTAGTTTTTTTGCAGCGCATAAAACATTCG
860 880 900 920 940 960
BglIIXhoII AluI BbvIFnu4HIAluIMboIIHphINcoI TaqI BbvIFnu4HIAluI
TATAGAATCTAAAATATCTAATCATGATCTTGGAGCTTCCTTGAAGGAGTGGCAGCTTCTTCACCCATGGCATGAATGATAGTGACCATCAAATCCTTACCTTCATCGAAGCAGCTTG
ATATCTTTAGATTTTATAGATTAGTCTAGAACCTCGAAGGAACTTCCTCTACCGTCGAAGAAGTGGGTACCGTCTTTACTATCACTGGTAGTTTAGGAATGGAAGTAGCTTTCGTCGAAC
980 1000 1020 1040 1060 1080
* D S R P A E K F S I A A E E G M A S I I T V M L D K G E D F A A Q
SphINsp(7524)I HphI HindIIITth111I HphI EcoRI* HinfIII
CATGCTGTCACCCAACCTACCTTCTGGTGCGTTGACATCGTCTTTGGTTTCACCGTTCATGGTCATCAAGGACAAGTAGCCATCATCATATATCCAACAATTGGTATTCGCTTCTGTTGAC
GTACGACAGTGGGTTGGATGGAAGACCACGCAACTGTAGCAGAAACCAAAGTGGCAAGTACCAGTAGTTCCTGTTCATCGGTAGTAGTTATAGGTTGTTAACCATAAGCGAAGAGAACTG
1100 1120 1140 1160 1180 1200
M S D G L R G E P A N V D D K T E G N M T M L S L Y G D D I D L L Q Y E S R K V
FokI MboII MboIIEcoRV BstEII AvaIISau96I EcoPIHpaII Nsp(7524)I
AAATGGAACTTCCAAGTTGTGAGTGGATGGAGACAAATCTTCCAACTTCTTATCCAGTGAAGATATCAAGGGTAACCAAATGGACGTTGGCGTGACCGTGCTTACCGGTCTTGGAAGTGGA
TTTACCTTGAAGGTTCAACACTCACCTACCTCTGTTTAGAAGGTTGAAGAATGGTCACTTCTATAGTTCCCATTGGTTTACCTGGAACCGCACTGGCACGAATGGCCAGAACCTTCACCT
1220 1240 1260 1280 1300 1320
F P V E L N H T S P S L D E L K K G T F I D L T V L H V K A H G H K G T K S T S


AccIIHindIISalITaqI AccIEcoPI XmnI MnlIHgiCI HgaI
1340 1360 1380 1400 1420 1440
M D V I K Q P R G K I V V F G N K R L A S C Q M
MboII Nsp(7524)I EcoRI* EcoRI* MboII
1460 1480 1500 1520 1540 1560
EcoRI* EcoRI* RsaI EcoRI* DdeI MboII
1580 1600 1620 1640 1660 1680
MboII EcoRI* XhoI
1700 1720 1740 1760 1780 1800
HphI MnlI EcoRI*
1820 1840 1860 1880 1900 1920
EcoRI* TthIIIITaqIHinfIII HinfI EcoRI* HinfI EcoRI*HpaIIScrFISmaI FnuEI
1940 1960 1980 2000 2020 2040
HphIFnuEI EcoPIHpaII CauIICfrIHaeIIIHpaIIScrFI MboIIEcoPI HaeIIIHhaIAluI BinIFnuEI FnuDII MnlIXhoI
2060 2080 2100 2120 2140 2160
FnuEIXhoIIBinI BstNIScrFI BalICfrIFokIHaeIIIBstNIScrFI BclI HgiAIAvaIIINsp(7524)ISphI
2180 2200 2220 2240 2260 2280
MboII AvaIISau96I EcoRI* TthIIII EcoRI*
2300 2320 2340 2360 2380 2400
M T E F K A G S A K K G A T L F K T R C L Q C H T V E K G G P H K V G P N L H
EcoRI HaeIIIHpaII DdeI AciI HaeIIISau96I AvaIISau96I
2420 2440 2460 2480 2500 2520
G I F G R H S G Q A E G Y S Y T D A N I K K N V L W D E N N M S E Y L T N P K K
AluI HinfIIIRsaI SfaNI Nsp(7524)I RsaISacI
2540 2560 2580 2600 2620 2640
Y I P G T K M A F G G L K K E K D R N D L I T Y L K K A C E *
BstNIScrFIKpnIRsaI HaeIII MboII EcoRI* HaeIIISau96I TaqIEcoRV


2660
2680
2700
2720
2740
2760
EcoRI*
MnlI
FokI
AvaIISau96I
2780
2800
2820
2840
2860
2880
EcoRI*
RsaI HgaIFnuDII RsaINsp(7524)I
TthIIIII
HgaI TaqI
EcoRI*
2900
2920
2940
2960
2980
3000
HindIII
3010

Appendix B

Experimental NOE Intensities for the Cyanide Adduct of *Saccharomyces cerevisiae* Ala80cytochrome *c*

Part I lists the experimental NOE intensities obtained from 2D NOESY and 1D NOE experiments used for the structural calculation. Part II lists irrelevant NOE constraints not used in the structural calculation. Part III lists hydrogen bond constraints used in the calculation of family DG3. An asterisk indicates peak intensities that correspond to two or more protons with degenerate chemical shifts; these intensities are divided by the number of protons they represent. CYSS denotes the Cys residues that are linked to the heme by thioether linkages. HES is an artificial amino acid including His18 and the heme. TML is trimethyllysine.

-3	PHE				HA	-2	LYS+	HB3	11844.0
HN	-3	PHE	HA	9596.0	HA	-2	LYS+	HG2	19471.2
HN	-3	PHE	HB2	20250.0	HA	-2	LYS+	HD2	14080.0
HN	-3	PHE	HB3	12840.0	HA	-2	LYS+	HE2	2524.0
HN	-3	PHE	CG	1664.5*2	HA	-1	ALA	HN	35930.0
HN	-2	LYS+	HN	670.0	HB2	-2	LYS+	HG2	48970.0
HN	-2	LYS+	HG2	2198.0	HB2	-2	LYS+	HE2	11230.8
HA	-3	PHE	HB2	53701.2	HB2	-1	ALA	HN	19630.0
HA	-3	PHE	HB3	12910.8	HB3	-2	LYS+	HG2	66553.2
HA	-3	PHE	CG	9479.4*2	HB3	-2	LYS+	HE2	11978.4
HA	-2	LYS+	HN	18380.0	HG2	-2	LYS+	HD2	103400.0
HB2	-3	PHE	CG	21425.0*2	HG2	-2	LYS+	HE2	8736.0
HB3	-3	PHE	CG	17180.0*2	HG2	-1	ALA	HN	7336.0
CG	-3	PHE	CZ	26450.0*2*2	HD2	-2	LYS+	HE2	10470.0
CG	-3	PHE	HZ	19560.0*2	-1	ALA			
CG	-2	LYS+	HN	4143.0*2	HN	-1	ALA	HA	13170.0
CG	-2	LYS+	HA	2374.0*2	HN	-1	ALA	QB	32650.0
CG	-1	ALA	HA	193.2*2	HA	-1	ALA	QB	80676.0
CG	61	GLU-	HB2	2371.7*2	HA	1	GLY	HN	18430.0
CG	61	GLU-	HB3	10655.0*2	HA	92	ASN	HB3	5803.6
CG	61	GLU-	HG2	2853.0*2	HA	96	THR	QG2	3847.2
CG	61	GLU-	HG3	1566.0*2	QB	1	GLY	HN	22790.0
CG	92	ASN	HA	4544.4*2	1	GLY			
CG	95	ILE	HB	3572.1*2	HN	1	GLY	HA1	2690.0
CG	95	ILE	QG2	5175.0*2	HN	1	GLY	HA2	23180.0
CG	95	ILE	QD1	3968.2*2	HN	92	ASN	HB3	8949.0
CZ	-3	PHE	HZ	135600.0*2	HN	96	THR	QG2	15240.0
CZ	-1	ALA	HA	5812.8*2	HA1	2	SER	HN	10840.0
CZ	-1	ALA	QB	8195.0*2	HA2	2	SER	HN	9317.0
CZ	61	GLU-	HB2	2863.6*2	HA2	93	ASP-	HA	3911.0
CZ	61	GLU-	HB3	9760.0*2	2	SER			
CZ	61	GLU-	HG2	1009.0*2	HN	2	SER	HB2	3049.0
CZ	61	GLU-	HG3	3126.0*2	HN	2	SER	HB3	5049.0
CZ	92	ASN	HA	10042.2*2	HN	5	LYS+	HB3	1316.0
CZ	92	ASN	HB3	5820.0*2	HN	93	ASP-	HA	6819.0
CZ	95	ILE	HA	1538.5*2	HA	2	SER	HB2	12566.4
CZ	95	ILE	HB	12877.2*2	HA	2	SER	HB3	17110.8
CZ	95	ILE	QG2	11430.0*2	HB3	5	LYS+	HB3	41916.0
CZ	95	ILE	QD1	3310.9*2	HB3	5	LYS+	HG2	7128.2
CZ	96	THR	HN	4533.5*2	3	ALA			
HZ	-1	ALA	HA	20000.4	HA	3	ALA	QB	42280.0
HZ	-1	ALA	QB	15860.0	HA	96	THR	QG2	5939.0
HZ	1	GLY	HN	5235.0	HA	97	TYR	HN	12000.0
HZ	61	GLU-	HG3	7021.0	QB	4	LYS+	HN	11560.0
HZ	92	ASN	HA	7817.0	QB	4	LYS+	HA	7047.0
HZ	95	ILE	HB	5753.2	4	LYS+			
HZ	95	ILE	QG2	10370.0	HN	4	LYS+	HA	13018.0
HZ	96	THR	HN	4265.0	HN	4	LYS+	HB2	11010.0
HZ	96	THR	HA	15649.2	HN	4	LYS+	HB3	4192.0
-2	LYS+				HN	4	LYS+	HE2	770.3
HN	-2	LYS+	HA	10300.0	HN	5	LYS+	HN	951.0
HN	-2	LYS+	HB2	5935.0	HA	4	LYS+	HB2	13860.0
HN	-2	LYS+	HG2	8921.0	HA	4	LYS+	HB3	34130.0
HN	-2	LYS+	HD2	14920.0	HA	5	LYS+	HN	4534.0
HN	-2	LYS+	HE2	2311.0	HA	6	GLY	HN	3254.0
HN	-1	ALA	HN	809.1	HA	7	ALA	HN	4190.0
HA	-2	LYS+	HB2	48207.6	HA	7	ALA	QB	9408.0

HB2	5	LYS+	HN	30630.0
HB3	5	LYS+	HN	4694.0
5 LYS+				
HN	5	LYS+	HA	40950.0
HN	5	LYS+	HB2	6839.0
HN	6	GLY	HN	7902.0
HA	5	LYS+	HB2	10730.0
HA	6	GLY	HN	3277.0
HA	8	THR	HN	2171.0
HB3	5	LYS+	HG2	31382.4
HB3	6	GLY	HN	19250.0
HB3	93	ASP-	HB3	35800.8
HG2	6	GLY	HN	1360.0
HG2	93	ASP-	HB3	32440.8
6 GLY				
HN	6	GLY	HA1	8684.0
HN	6	GLY	HA2	17690.0
HN	7	ALA	HN	13730.0
HN	93	ASP-	HB3	10760.0
HN	97	TYR	HB2	1834.0
HA1	7	ALA	HN	5130.0
HA1	9	LEU	QD2	29761.2
HA1	94	LEU	HA	25586.4
HA2	7	ALA	HN	9481.0
HA2	94	LEU	HA	18463.2
7 ALA				
HN	7	ALA	HA	20280.0
HN	7	ALA	QB	75340.0
HN	8	THR	HN	23970.0
HN	97	TYR	HB2	10794.0
HA	7	ALA	QB	100632.0
HA	10	PHE	HN	4661.0
HA	10	PHE	HB2	7481.9
HA	10	PHE	HB3	29895.6
QB	8	THR	HN	26340.0
8 THR				
HN	8	THR	HA	14010.0
HN	8	THR	HB	52890.0
HN	8	THR	QG2	6984.0
HN	9	LEU	HN	42090.0
HN	11	LYS+	HB2	3350.0
HA	8	THR	HB	12280.8
HA	8	THR	QG2	32790.0
HA	9	LEU	HN	5639.0
HA	11	LYS+	HN	5830.0
HA	11	LYS+	HB2	38547.6
HA	11	LYS+	HB3	7011.0
HB	8	THR	QG2	57766.8
HB	9	LEU	HN	2223.0
HB	12	THR	QG2	4249.0
QG2	9	LEU	HN	6203.0
9 LEU				
HN	9	LEU	HA	19900.0
HN	9	LEU	HB2	11090.0
HN	9	LEU	HB3	19310.0
HN	9	LEU	QD1	11390.0
HN	9	LEU	QD2	7671.0
HN	10	PHE	HN	18760.0
HN	11	LYS+	HN	6684.0
HN	94	LEU	QD1	1440.0
HN	94	LEU	QD2	2442.0
HA	9	LEU	HB2	9559.2
HA	9	LEU	HB3	22990.8
HA	9	LEU	QD1	53496.0
HA	9	LEU	QD2	42048.0
HA	10	PHE	HN	5005.0
HA	13	ARG+	HN	1782.0
HA	13	ARG+	HG2	7644.8
HB2	9	LEU	QD1	18396.0
HB2	9	LEU	QD2	169344.0
HB2	10	PHE	HN	15610.0
HB2	94	LEU	HB3	30470.0
HB2	94	LEU	QD1	7350.0
HB2	94	LEU	QD2	12096.0
HB3	9	LEU	QD1	51460.0
HB3	9	LEU	QD2	18312.0
HB3	10	PHE	HN	1863.0
QD1	10	PHE	HN	555.0
QD1	13	ARG+	HN	1355.0
QD1	13	ARG+	HG2	15450.0
QD1	90	ASP-	HB3	1338.1
QD2	10	PHE	HN	11700.0
QD2	85	LEU	QD1	17942.4
QD2	90	ASP-	HN	750.5
QD2	90	ASP-	HA	57355.2
QD2	90	ASP-	HB3	4515.8
QD2	94	LEU	HN	15250.0
QD2	94	LEU	HB2	46746.0
QD2	94	LEU	HB3	19345.2
QD2	94	LEU	QD1	6456.2
QD2	94	LEU	QD2	7312.2
10 PHE				
HN	10	PHE	HA	20760.0
HN	10	PHE	HB2	40760.0
HN	10	PHE	HB3	25800.0
HN	10	PHE	HD2	1371.0
HN	11	LYS+	HN	18740.0
HN	94	LEU	QD1	8063.0
HN	94	LEU	QD2	11210.0
HA	10	PHE	HB2	26787.6
HA	10	PHE	HB3	16531.2
HA	10	PHE	HD2	20958.0
HA	11	LYS+	HN	4408.0
HA	13	ARG+	HN	2366.0
HA	14	CYSS	HN	7285.0
HA	14	CYSS	HB2	10510.0
HA	94	LEU	QD1	18396.0
HA	94	LEU	QD2	21820.0
HB2	10	PHE	HD2	32986.8
HB2	11	LYS+	HN	3526.0
HB2	94	LEU	QD1	6838.4
HB2	94	LEU	QD2	6569.6
HB3	10	PHE	HD2	21151.2
HB3	11	LYS+	HN	30560.0
HE1	10	PHE	HD2	13498.8
HE1	20	VAL	QG2	4918.0
HE1	98	LEU	QD2	525.8
HZ	10	PHE	HE2	5597.0
HZ	18	HES	HB3	1437.0
HE2	10	PHE	HD2	12800.0
HE2	14	CYSS	HB2	38390.0
HE2	18	HES	QM1	3408.0
HE2	19	THR	HA	4932.0
HE2	20	VAL	QG1	28030.0
HE2	32	LEU	QD1	15350.0
HE2	98	LEU	QD2	2690.0
HD2	14	CYSS	HB2	12950.0
HD2	18	HES	QM1	12060.0
HD2	94	LEU	QD1	12348.0
HD2	94	LEU	QD2	22280.0
HD2	98	LEU	QD1	10780.0
HD2	98	LEU	QD2	9769.2
11 LYS+				
HN	11	LYS+	HA	11360.0
HN	11	LYS+	HB2	52110.0
HN	11	LYS+	HB3	5633.0

HN	12	THR	HN	18240.0
HA	11	LYS+	HB2	8821.0
HA	11	LYS+	HB3	22117.2
HA	11	LYS+	HE2	15590.4
HA	12	THR	HN	5029.0
HA	15	LEU	HN	11600.0
HA	15	LEU	HA	3075.0
HA	15	LEU	HB2	14305.2
HA	15	LEU	QD2	29164.8
HB2	11	LYS+	HE2	24880.0
HB2	12	THR	HN	15040.0
HB2	12	THR	QG2	14733.6
HB3	11	LYS+	HE2	7070.0
HB3	12	THR	HN	3951.0
HB3	12	THR	QG2	6021.1
12	THR			
HN	12	THR	HA	7126.0
HN	12	THR	QG2	23910.0
HN	13	ARG+	HN	17340.0
HA	12	THR	QG2	39505.2
HA	13	ARG+	HN	2250.0
QG2	13	ARG+	HN	1718.0
13	ARG+			
HN	13	ARG+	HA	5015.1
HN	13	ARG+	HB2	17950.0
HN	13	ARG+	HB3	3620.0
HN	13	ARG+	HG2	14470.0
HN	14	CYSS	HN	65310.0
HN	14	CYSS	HA	798.0
HN	15	LEU	HN	12320.0
HA	13	ARG+	HB2	2025.0
HA	13	ARG+	HB3	27854.4
HB2	14	CYSS	HN	8322.0
HB2	18	HES	HT2A	3098.0
HB2	18	HES	QT2	6202.0
HB3	14	CYSS	HN	4833.0
HB3	18	HES	HT2A	6100.0
HB3	18	HES	QT2	13900.0
HB3	82	PHE	HZ	5411.3
HB3	82	PHE	QR	2289.8*5
14	CYSS			
HN	14	CYSS	HA	10340.0
HN	14	CYSS	HB2	3072.0
HN	15	LEU	HA	593.0
HN	15	LEU	HB2	1268.0
HN	15	LEU	HB3	1432.0
HN	16	GLN	HN	2269.0
HA	14	CYSS	HB2	3370.0
HA	18	HES	HT2A	38080.0
HA	18	HES	HAM	17780.0
HA	18	HES	QM3	16270.0
HB2	18	HES	HB3	1557.0
HB2	18	HES	QM1	2796.0
15	LEU			
HN	15	LEU	HA	20670.0
HN	15	LEU	HB2	32960.0
HN	15	LEU	HB3	37820.0
HN	15	LEU	QD2	12830.0
HN	16	GLN	HN	9088.0
HN	17	CYSS	HN	2433.0
HA	15	LEU	HB2	37833.6
HA	15	LEU	HB3	14523.6
HA	15	LEU	HG	20412.0
HA	15	LEU	QD1	6926.6
HA	15	LEU	QD2	18480.0
HA	16	GLN	HN	2890.0
HA	17	CYSS	HN	2377.0
HA	18	HES	HN	13400.0
HB2	15	LEU	HG	52640.0
HB2	15	LEU	QD1	108780.0
HB2	16	GLN	HN	2260.0
HB3	15	LEU	QD1	25998.0
HB3	15	LEU	QD2	55090.0
HB3	16	GLN	HN	7937.0
HG	15	LEU	QD1	38673.6
HG	15	LEU	QD2	112400.0
HG	16	GLN	HN	13580.0
QD1	16	GLN	HN	2881.0
QD1	19	THR	HA	2499.0
16	GLN			
HN	16	GLN	HA	3245.0
HN	16	GLN	HB2	19500.0
HN	16	GLN	HG2	13250.0
HN	16	GLN	HG3	2760.0
HN	17	CYSS	HN	14110.0
HN	18	HES	HN	3222.0
HN	18	HES	QM3	618.0
HA	16	GLN	HB3	20538.0
HA	16	GLN	HG2	13818.0
HA	16	GLN	HG3	13482.0
HB2	16	GLN	HE21	1806.0
HB2	18	HES	QM3	5136.0
HB3	16	GLN	HG3	110500.0
HG2	16	GLN	HE21	1140.0
HG2	16	GLN	HE22	9081.0
HG2	17	CYSS	HN	6611.0
HG3	16	GLN	HE21	681.0
HG3	18	HES	QT4	9777.0
HE21	18	HES	QT4	890.0
HE22	18	HES	QT4	2762.0
17	CYSS			
HN	17	CYSS	HA	14640.0
HN	18	HES	HN	64970.0
HN	18	HES	QM3	5832.0
HA	17	CYSS	HB2	2559.5
HA	17	CYSS	HB3	15860.0
HA	18	HES	HN	4802.0
HA	27	LYS+	HB2	5189.5
HA	27	LYS+	HB3	5749.0
HA	27	LYS+	HE3	3725.4
HA	28	VAL	HN	7800.0
HA	28	VAL	HB	54690.0
HA	28	VAL	QG1	8451.0
HA	28	VAL	QG2	17190.0
HA	29	GLY	HN	12090.0
HB3	18	HES	HN	5949.0
HB3	18	HES	HT4A	10010.0
HB3	28	VAL	HB	2903.0
HB3	28	VAL	QG1	1885.0
HB3	29	GLY	HN	9958.0
18	HES			
HN	18	HES	HA	14540.0
HN	18	HES	HB2	20000.0
HN	18	HES	HB3	3213.0
HN	19	THR	HN	1064.0
HN	19	THR	QG2	2419.0
HA	18	HES	HB2	12410.0
HA	18	HES	HB3	25110.0
HA	19	THR	HN	60310.0
HA	19	THR	HA	1314.0
HA	27	LYS+	HB2	1169.0
HA	31	ASN	HA	7524.0
HB2	19	THR	HN	8217.0
HB2	32	LEU	QD1	8282.0
HB3	19	THR	HN	16940.0
HB3	19	THR	HA	677.0

HB3	32	LEU	HN	1352.0
HB3	32	LEU	HG	10770.0
HB3	32	LEU	QD1	13350.0
HB3	32	LEU	QD2	.0
HD1	32	LEU	HG	7586.0
HAP71	18	HES	HAP7	252000.0
HAP71	18	HES	HBP7	18975.6
HAP71	18	HES	HBP7	55000.0
HAP71	18	HES	QM8	10500.0
HAP71	52	ASN	HD21	9766.0
HAP71	59	TRP	HE1	1214.0
HAP71	59	TRP	HZ2	20521.2
HAP72	18	HES	HBP7	63000.0
HAP72	18	HES	HBP7	45360.0
HAP72	18	HES	QM8	11000.0
HAP72	52	ASN	HD21	3031.0
HAP72	59	TRP	HE1	7110.0
HAP72	59	TRP	HZ2	33180.0
HAP72	59	TRP	HH2	11140.0
HAP72	67	TYR	HE1	1837.0
HBP73	18	HES	QM8	6682.0
HBP73	32	LEU	QD2	3083.6
HBP73	35	ILE	QD1	17707.2
HBP73	48	TYR	HH	1896.3
HBP73	52	ASN	HD22	2547.2
HBP73	59	TRP	HE1	744.0
HBP73	59	TRP	HZ2	1326.0
HBP74	30	PRO	HG3	9735.0
HBP74	30	PRO	HD2	6465.0
HBP74	48	TYR	HE1	1850.0
HBP74	48	TYR	HH	3392.5
HBP74	52	ASN	HD22	3882.4
QM8	18	HES	HDM	25180.0
QM8	32	LEU	QD1	3501.0
QM8	32	LEU	QD2	7661.0
QM8	35	ILE	QG2	2879.0
QM8	35	ILE	HG12	10880.0
QM8	35	ILE	HG13	3693.0
QM8	35	ILE	QD1	17870.0
QM8	59	TRP	HE1	1741.0
QM8	59	TRP	HZ2	3195.0
QM8	59	TRP	HH2	2648.0
QM8	64	MET	QE	29830.0
QM8	68	LEU	QD1	1212.0
HDM	18	HES	QM1	35170.0
HDM	32	LEU	QD1	6624.0
HDM	32	LEU	QD2	5667.0
HDM	68	LEU	QD1	7208.0
HDM	68	LEU	QD2	3944.0
HDM	98	LEU	QD1	1603.0
QM1	18	HES	QT2	8280.0
QM1	32	LEU	QD1	22300.0
QM1	32	LEU	QD2	3558.0
QM1	68	LEU	HG	2200.0
QM1	68	LEU	QD1	37360.0
QM1	68	LEU	QD2	48390.0
QM1	94	LEU	QD1	49960.0
QM1	94	LEU	QD2	16120.0
QM1	98	LEU	QD1	38040.0
QM1	98	LEU	QD2	20670.0
HT2A	18	HES	QT2	114000.0
HT2A	18	HES	HAM	24000.0
HT2A	18	HES	QM3	6906.0
QT2	68	LEU	QD1	18110.0
QT2	68	LEU	QD2	32090.0
QT2	82	PHE	HZ	30270.0
QT2	82	PHE	QR	8522.0*5
QT2	85	LEU	QD1	16312.8
QT2	85	LEU	QD2	12658.8
QT2	94	LEU	QD1	8103.5
QT2	94	LEU	QD2	24318.0
HAM	18	HES	QM3	25400.0
HAM	82	PHE	HB2	10280.0
QM3	18	HES	HT4A	7000.0
QM3	18	HES	QT4	21210.0
QM3	82	PHE	HB2	10400.0
QM3	82	PHE	QR	894.8*5
HT4A	18	HES	QT4	68990.0
HT4A	18	HES	QM5	8045.0
HT4A	28	VAL	QG1	75790.0
HT4A	80	ALA	HA	1084.0
HT4A	81	ALA	HN	6014.0
QT4	28	VAL	QG1	12400.0
QT4	81	ALA	HN	9630.0
HBM	18	HES	QM5	39800.0
HBM	80	ALA	HA	6118.0
HBM	81	ALA	HN	3881.0
QM5	18	HES	HAP6	11410.0
QM5	18	HES	HAP6	30530.0
QM5	18	HES	HBP6	20991.6
QM5	28	VAL	HA	1295.3
QM5	28	VAL	QG1	36300.0
QM5	28	VAL	QG2	3803.0
QM5	29	GLY	HN	2963.0
QM5	29	GLY	HA1	13370.0
QM5	46	TYR	HE2	2595.0
QM5	79	LYS+	HN	5268.0
QM5	79	LYS+	HB3	13339.2
QM5	79	LYS+	HG2	6087.0
QM5	79	LYS+	HG3	19540.0
QM5	79	LYS+	HD2	6543.0
QM5	79	LYS+	HE3	16110.0
QM5	80	ALA	HN	4411.0
QM5	80	ALA	HA	21290.0
QM5	81	ALA	HN	4358.0
HAP61	18	HES	HBP6	17440.0
HAP61	46	TYR	HE2	10820.0
HAP62	18	HES	HBP6	59380.0
HAP62	46	TYR	HE2	11670.0
HBP63	79	LYS+	HN	13410.0
HBP63	80	ALA	HN	13540.0
HBP63	80	ALA	HA	1798.0
HBP63	80	ALA	QB	83210.4

19 THR

HN	19	THR	HA	14900.0
HN	19	THR	HB	3956.0
HN	19	THR	QG2	13740.0
HN	20	VAL	HN	672.8
HN	27	LYS+	HG3	1603.0
HN	31	ASN	HA	17440.0
HN	32	LEU	HN	12530.0
HA	19	THR	HB	42142.8
HA	19	THR	QG2	47544.0
HA	20	VAL	HN	51290.0
HA	20	VAL	QG1	7843.1
HA	21	GLU-	HN	4775.0
HB	19	THR	QG2	65721.6
HB	20	VAL	HN	23230.0
HB	20	VAL	HB	1709.4
HB	20	VAL	QG1	858.0
HB	21	GLU-	HN	11800.0
HB	24	GLY	HA2	8351.3
HB	31	ASN	HD21	4753.0
HB	31	ASN	HD22	3076.0
QG2	27	LYS+	HE2	23066.4

20 VAL

HN	20	VAL	HB	22977.0
HN	20	VAL	QG1	43953.0
HN	20	VAL	QG2	8038.5
HN	21	GLU-	HN	43447.0
HN	21	GLU-	HG2	13271.0
HA	20	VAL	HB	22226.4
HA	20	VAL	QG1	8433.6
HA	20	VAL	QG2	14540.4
HB	20	VAL	QG1	37674.0
HB	20	VAL	QG2	60866.4
HB	98	LEU	QD2	82.9
QG1	21	GLU-	HN	22057.0
QG1	32	LEU	QD1	2859.0
QG1	101	ALA	QB	9089.6
QG2	98	LEU	QD1	2160.5
QG2	98	LEU	QD2	21620.0
QG2	102	SER	HN	3585.0
21	GLU-			
HN	21	GLU-	HG2	29520.0
HN	21	GLU-	HG3	28130.0
HN	24	GLY	HA2	1043.0
HN	31	ASN	HD22	2985.0
HA	21	GLU-	HB2	20874.0
HA	21	GLU-	HB3	14246.4
HA	21	GLU-	HG2	2459.5
HA	21	GLU-	HG3	4717.4
HG3	22	LYS+	HN	6145.0
22	LYS+			
HN	22	LYS+	HA	18740.0
HN	22	LYS+	HB2	15960.0
HN	22	LYS+	HB3	8539.0
HN	22	LYS+	HG2	2630.0
HN	22	LYS+	HG3	10720.0
HA	22	LYS+	HB2	8467.2
HA	22	LYS+	HB3	24561.6
HA	22	LYS+	HG2	46695.6
HA	22	LYS+	HG3	39190.0
HA	22	LYS+	HD2	19219.2
HA	23	GLY	HN	20110.0
HA	24	GLY	HN	1823.0
HA	33	HIS	HD2	4520.9
HB2	22	LYS+	HG2	43671.6
HB2	22	LYS+	HD2	14599.2
HB2	22	LYS+	HE3	29727.6
HB3	22	LYS+	HG3	42714.0
HB3	22	LYS+	HD2	37783.2
HB3	22	LYS+	HE3	32650.8
HB3	23	GLY	HN	3535.0
HB3	33	HIS	HD2	1438.0
HB3	33	HIS	HE1	4701.0
HG2	22	LYS+	HD2	36036.0
HG2	22	LYS+	HE3	6226.1
HG2	23	GLY	HN	5500.0
HG2	33	HIS	HD2	2424.0
HG2	33	HIS	HE1	603.0
HG3	22	LYS+	HD2	23562.0
HG3	23	GLY	HN	2140.0
HG3	33	HIS	HD2	2454.0
HG3	33	HIS	HE1	2823.0
HD2	22	LYS+	HE3	19597.2
HD2	23	GLY	HN	3413.0
HD2	33	HIS	HE1	1575.0
23	GLY			
HN	23	GLY	HA2	346.0
HN	24	GLY	HN	3776.0
HN	24	GLY	HA2	700.0
HA1	24	GLY	HN	5392.4
HA2	24	GLY	HN	10972.1
24	GLY			
HN	24	GLY	HA1	3806.5
HN	24	GLY	HA2	7357.0
HN	31	ASN	HD21	3895.0
HN	31	ASN	HD22	9382.0
25	PRO			
HA	25	PRO	HB3	32365.2
HA	26	HIS	HN	3214.0
HB2	26	HIS	HN	2722.0
HB3	25	PRO	HG2	241416.0
HB3	25	PRO	HG3	21831.6
HB3	26	HIS	HN	4261.0
HG3	25	PRO	HD2	104496.0
26	HIS			
HN	26	HIS	HA	3817.0
HN	26	HIS	HB2	4043.0
HN	26	HIS	HB3	1534.0
HA	26	HIS	HB2	13734.0
HA	26	HIS	HB3	17925.6
HA	26	HIS	HD2	3925.3
HA	27	LYS+	HN	24070.0
HA	27	LYS+	HG3	11121.6
HA	30	PRO	HA	13658.4
HA	31	ASN	HN	11060.0
HA	31	ASN	HB2	6388.2
HA	31	ASN	HB3	4431.8
HB2	26	HIS	HD2	11121.6
HB2	30	PRO	HA	6949.0
HB2	31	ASN	HN	488.0
HB2	46	TYR	HD1	11558.4
HB2	46	TYR	HE1	561.6
HB2	46	TYR	HE2	1712.0
HB3	26	HIS	HD2	16682.4
HB3	30	PRO	HA	11667.6
HB3	46	TYR	HD1	13431.6
HB3	46	TYR	HE2	7374.0
HB3	46	TYR	HD2	4419.2
HD2	30	PRO	HA	854.0
HD2	46	TYR	HA	4207.6
HD2	46	TYR	HB2	3653.2
HE1	30	PRO	HA	3212.2
HE1	30	PRO	HB2	4332.0
HE1	30	PRO	HB3	3581.8
HE1	31	ASN	HN	17100.0
HE1	31	ASN	HB2	9357.6
HE1	31	ASN	HB3	19244.4
HE1	43	ALA	QB	27200.0
HE1	46	TYR	HB2	3780.0
HE1	46	TYR	HB3	6108.0
27	LYS+			
HN	27	LYS+	HA	5593.0
HN	27	LYS+	HB2	9559.0
HN	27	LYS+	HB3	7055.0
HN	27	LYS+	HG3	9619.0
HN	27	LYS+	HD2	1409.0
HN	27	LYS+	HE2	8233.0
HN	27	LYS+	HE3	4177.0
HN	28	VAL	HN	1401.0
HN	29	GLY	HN	5065.0
HN	30	PRO	HA	723.0
HN	31	ASN	HN	993.0
HA	27	LYS+	HB2	13473.6
HA	27	LYS+	HB3	34666.8
HA	27	LYS+	HG3	42268.8
HA	27	LYS+	HD2	12100.0
HA	27	LYS+	HD3	6362.0
HA	27	LYS+	HE2	4860.0
HA	27	LYS+	HE3	5287.0

HA	28	VAL	HN	1954.0
HA	28	VAL	QG2	3187.8
HB2	27	LYS+	HG3	14288.4
HB2	27	LYS+	HD2	24612.0
HB2	27	LYS+	HE3	23930.0
HB2	28	VAL	HN	22020.0
HB3	27	LYS+	HG3	33507.6
HB3	27	LYS+	HD2	75936.0
HB3	27	LYS+	HE2	4204.2
HB3	27	LYS+	HE3	3667.4
HB3	28	VAL	HN	35900.0
HG3	27	LYS+	HD2	40404.0
HG3	28	VAL	HN	6796.0
HD2	27	LYS+	HE2	6650.3
HD2	27	LYS+	HE3	9075.0
HD3	27	LYS+	HE2	565.0
28 VAL				
HN	28	VAL	HA	15090.0
HN	28	VAL	HB	29190.0
HN	28	VAL	QG1	8189.0
HN	28	VAL	QG2	32870.0
HN	29	GLY	HN	17880.0
HA	28	VAL	HB	11630.0
HA	28	VAL	QG1	47200.0
HA	28	VAL	QG2	52810.8
HA	29	GLY	HN	4583.0
HB	28	VAL	QG1	49490.0
HB	28	VAL	QG2	46905.6
HB	29	GLY	HN	15200.0
HB	29	GLY	HA2	1344.0
QG1	29	GLY	HN	7532.0
29 GLY				
HN	29	GLY	HA1	2281.0
30 PRO				
HA	30	PRO	HB2	52819.2
HA	30	PRO	HB3	37279.2
HA	30	PRO	HG2	3008.0
HA	30	PRO	HG3	3188.0
HA	30	PRO	HD2	1256.0
HA	31	ASN	HN	32680.0
HA	46	TYR	HD1	8887.2
HB2	30	PRO	HG3	3667.0
HB2	31	ASN	HN	5614.0
HB2	43	ALA	QB	17908.8
HB3	30	PRO	HG3	39730.0
HB3	31	ASN	HN	9466.0
HB3	43	ALA	QB	31441.2
HG2	32	LEU	QD2	13920.0
HG2	48	TYR	HH	4083.6
HG3	30	PRO	HD2	9034.0
31 ASN				
HN	31	ASN	HA	12250.0
HN	31	ASN	HB2	18020.0
HN	31	ASN	HB3	23620.0
HN	32	LEU	HN	759.0
HN	43	ALA	QB	3152.0
HA	31	ASN	HB2	36909.6
HA	31	ASN	HB3	12440.0
HA	31	ASN	HD21	9748.0
HA	31	ASN	HD22	5130.0
HA	32	LEU	HN	53660.0
HB2	31	ASN	HD21	14480.0
HB2	31	ASN	HD22	8227.0
HB2	32	LEU	HN	1508.0
HB3	31	ASN	HD21	9635.0
HB3	31	ASN	HD22	2554.0
HD21	32	LEU	HN	8230.0
HD21	33	HIS	HN	291.2
HD22	32	LEU	HN	2790.0
32 LEU				
HN	32	LEU	HA	1862.0
HN	32	LEU	HB2	27550.0
HN	32	LEU	HB3	5234.0
HN	32	LEU	HG	10410.0
HN	32	LEU	QD1	6689.0
HN	32	LEU	QD2	4399.0
HN	33	HIS	HN	14310.0
HA	32	LEU	HB2	10785.6
HA	32	LEU	HB3	28442.4
HA	32	LEU	QD1	5304.6
HA	32	LEU	QD2	41286.0
HA	35	ILE	HB	18320.4
HA	35	ILE	QG2	6985.4
HA	35	ILE	QD1	5684.3
HB2	32	LEU	HG	21117.6
HB2	32	LEU	QD1	20580.0
HB2	32	LEU	QD2	1128.1
HB2	35	ILE	QD1	424.0
HB2	98	LEU	QD2	10040.0
HB3	32	LEU	HG	17800.0
HB3	32	LEU	QD1	52790.0
HB3	32	LEU	QD2	19168.8
HB3	98	LEU	QD1	9724.0
HB3	98	LEU	QD2	17280.0
HG	32	LEU	QD1	21050.0
HG	32	LEU	QD2	20361.6
HG	35	ILE	QG2	3976.6
QD1	98	LEU	QD1	19610.0
QD1	98	LEU	QD2	12500.0
QD2	35	ILE	HB	9413.0
QD2	35	ILE	HG12	49600.0
33 HIS				
HN	33	HIS	HA	11670.0
HN	33	HIS	HB2	13310.0
HN	33	HIS	HB3	10070.0
HA	33	HIS	HB2	10730.0
HA	33	HIS	HB3	19850.0
HA	34	GLY	HN	13080.0
HA	102	SER	HA	28845.6
HB2	33	HIS	HD2	6951.0
HB3	33	HIS	HD2	8601.6
HD2	34	GLY	HA1	502.3
34 GLY				
HN	34	GLY	HA1	25200.0
HN	34	GLY	HA2	313.0
HN	35	ILE	HN	10330.0
HN	103	GLU-	HA	2430.0
HA1	35	ILE	HN	12790.0
35 ILE				
HN	35	ILE	HA	10370.0
HN	35	ILE	HB	47260.0
HN	35	ILE	QG2	33360.0
HN	35	ILE	HG12	5763.0
HN	35	ILE	HG13	7126.0
HN	35	ILE	QD1	7016.0
HA	35	ILE	HB	42820.0
HA	35	ILE	QG2	22110.0
HA	35	ILE	HG12	12583.2
HA	35	ILE	HG13	24750.0
HA	35	ILE	QD1	61420.0
HA	36	PHE	HN	2664.0
HA	38	ARG+	HG2	14666.4
HA	59	TRP	HD1	4175.6
HB	35	ILE	QG2	56926.8
HB	35	ILE	HG12	40610.0
HB	35	ILE	HG13	25640.0

HB	35	ILE	QD1	49310.0
HB	59	TRP	HD1	555.2
QG2	35	ILE	HG12	10290.0
QG2	35	ILE	HG13	28070.0
QG2	35	ILE	QD1	57370.0
QG2	36	PHE	HN	39860.0
QG2	36	PHE	HA	5658.2
QG2	36	PHE	CG	3769.5*2
QG2	36	PHE	CZ	7265.0*2
QG2	36	PHE	HZ	5072.0
QG2	59	TRP	HD1	759.4
QG2	102	SER	HB2	9886.8
HG12	35	ILE	QD1	42290.0
HG12	59	TRP	HD1	5177.0
HG13	35	ILE	QD1	81120.0
HG13	59	TRP	HD1	11660.0
QD1	59	TRP	HD1	36930.0
QD1	59	TRP	HE1	16190.0
QD1	59	TRP	HZ2	3077.0
QD1	64	MET	QE	13030.0
36 PHE				
HN	36	PHE	HA	11360.0
HN	36	PHE	HB2	29150.0
HN	36	PHE	HB3	4465.0
HN	37	GLY	HN	2199.0
HA	36	PHE	HB2	16090.0
HA	36	PHE	HB3	17967.6
HA	36	PHE	CG	8235.0*2
HA	36	PHE	CZ	1019.0*2
HA	37	GLY	HN	22600.0
HA	38	ARG+	HN	3107.0
HB2	36	PHE	CG	13395.0*2
HB3	36	PHE	CG	12140.0*2
CG	36	PHE	CZ	21504.0*2*2
CG	36	PHE	HZ	29370.0*2
CG	37	GLY	HN	1863.5*2
CG	60	ASP-	HA	9966.6*2
CG	61	GLU-	HN	964.5*2
CG	64	MET	QE	2427.0*2
CG	98	LEU	HB2	2544.8*2
CG	98	LEU	QD1	1308.7*2
CG	99	LYS+	HA	10823.4*2
CG	102	SER	HB2	9105.0*2
CZ	36	PHE	HZ	39719.4*2
CZ	60	ASP-	HA	1270.8*2
CZ	61	GLU-	HN	1638.5*2
CZ	61	GLU-	HA	13830.0*2
CZ	61	GLU-	HB3	4504.0*2
CZ	64	MET	HB2	2303.5*2
CZ	64	MET	HB3	1367.0*2
CZ	64	MET	QE	8980.0*2
CZ	95	ILE	HA	2135.7*2
CZ	95	ILE	QG2	9945.0*2
CZ	98	LEU	HA	1191.1*2
CZ	98	LEU	HB2	20640.0*2
CZ	98	LEU	HB3	12560.0*2
CZ	98	LEU	HG	3690.5*2
CZ	98	LEU	QD1	8590.0*2
CZ	98	LEU	QD2	1674.5*2
CZ	99	LYS+	HN	2887.5*2
CZ	99	LYS+	HA	12028.8*2
HZ	61	GLU-	HA	5367.0
HZ	64	MET	QE	14530.0
HZ	95	ILE	HA	7152.6
HZ	95	ILE	QG2	27680.0
HZ	95	ILE	HG12	5561.0
HZ	98	LEU	HB2	15370.0
HZ	98	LEU	HB3	7898.0
HZ	98	LEU	HG	3110.0
HZ	98	LEU	QD1	16820.0
HZ	98	LEU	QD2	6273.1
HZ	99	LYS+	HN	1277.0
HZ	99	LYS+	HA	2823.2
37 GLY				
HN	37	GLY	HA1	8438.0
HN	37	GLY	HA2	9992.0
HN	38	ARG+	HN	35340.0
HN	58	LEU	QD1	1847.0
HA1	58	LEU	QD1	15657.6
HA2	38	ARG+	HN	1657.0
38 ARG+				
HN	38	ARG+	HB2	5663.0
HN	38	ARG+	HB3	15940.0
HN	38	ARG+	HG2	5879.0
HN	58	LEU	QD1	3618.0
HA	38	ARG+	HB3	4493.2
HA	38	ARG+	HG2	6436.9
HB2	38	ARG+	HG2	152300.0
HG2	59	TRP	HD1	2646.8
39 GLN				
HN	39	GLN	HA	7207.0
HN	39	GLN	HB2	101000.0
HN	39	GLN	HB3	7622.0
HN	42	GLN	HG3	8094.0
HN	59	TRP	HD1	3322.0
HA	39	GLN	HB2	49946.4
HA	39	GLN	HB3	14187.6
HA	40	SER	HN	13190.0
HA	58	LEU	HA	82454.4
HA	58	LEU	HB2	8643.6
HA	58	LEU	HB3	6131.2
HA	58	LEU	QD1	6862.8
HA	59	TRP	HD1	2358.7
HB2	40	SER	HN	27010.0
HB2	56	ASN	HD21	1932.0
HB3	40	SER	HN	4803.0
40 SER				
HN	40	SER	HA	3398.2
HN	40	SER	HB2	5873.0
HN	40	SER	HB3	8424.0
HN	56	ASN	HA	758.0
HN	57	VAL	HN	7510.0
HN	57	VAL	HB	9509.0
HN	57	VAL	QG1	6981.0
HN	57	VAL	QG2	3424.0
HN	58	LEU	HA	2125.0
HN	59	TRP	HD1	653.0
HA	40	SER	HB2	26317.2
HA	40	SER	HB3	51618.0
HA	41	GLY	HN	18630.0
HA	59	TRP	HD1	7181.2
HA	59	TRP	HE1	23759.0
HA	59	TRP	HZ2	1337.0
HB2	41	GLY	HN	3183.0
HB2	57	VAL	QG1	10004.4
HB2	57	VAL	QG2	4399.1
HB2	59	TRP	HZ2	8259.7
HB3	41	GLY	HN	225.0
HB3	57	VAL	HB	17455.2
HB3	57	VAL	QG1	19488.0
HB3	57	VAL	QG2	9130.8
HB3	59	TRP	HE1	2152.0
HB3	59	TRP	HZ2	12020.4
41 GLY				
HN	52	ASN	HD22	2953.0
HN	59	TRP	HE1	12535.0

42	GLN			
HN	42	GLN	HA	1938.9
HN	42	GLN	HB2	17360.0
HN	42	GLN	HB3	3070.0
HN	42	GLN	HG2	10710.0
HN	42	GLN	HG3	30450.0
HN	43	ALA	HN	17600.0
HN	48	TYR	HE2	8151.0
HA	42	GLN	HB2	34700.4
HA	42	GLN	HB3	7559.0
HA	42	GLN	HG2	14179.2
HA	42	GLN	HG3	31340.4
HA	43	ALA	HN	2630.0
HB2	42	GLN	HG2	42495.6
HB2	42	GLN	HG3	28198.8
HB2	43	ALA	HN	4353.0
HB3	42	GLN	HG2	38472.0
HB3	42	GLN	HG3	18463.2
HG2	42	GLN	HE21	22480.0
HG2	42	GLN	HE22	13700.0
HG3	42	GLN	HE21	7774.0
HG3	42	GLN	HE22	2549.0
43	ALA			
HN	43	ALA	HA	1589.0
HN	43	ALA	QB	35570.0
HN	48	TYR	HD1	2505.0
HN	48	TYR	HE1	1695.0
HN	48	TYR	HE2	3981.0
HN	48	TYR	HH	7468.1
HA	43	ALA	QB	111216.0
HA	44	GLU-	HN	28300.0
QB	44	GLU-	HN	8646.0
QB	48	TYR	HE1	2403.2
QB	48	TYR	HH	7969.5
44	GLU-			
HN	44	GLU-	HA	7128.0
HN	44	GLU-	HB2	67010.0
HN	44	GLU-	HG2	6472.0
HN	44	GLU-	HG3	6989.0
HA	44	GLU-	HB2	64545.6
HA	44	GLU-	HG2	14170.8
HA	44	GLU-	HG3	11533.2
HA	45	GLY	HN	2959.0
HA	46	TYR	HN	2622.0
HB2	45	GLY	HN	5758.0
45	GLY			
HN	45	GLY	HA1	4147.0
HN	45	GLY	HA2	3132.0
HA1	46	TYR	HN	4343.0
HA2	46	TYR	HN	6699.0
46	TYR			
HN	46	TYR	HA	1012.2
HA	46	TYR	HB2	42294.0
HA	46	TYR	HB3	13140.0
HA	46	TYR	HD1	16212.0
HA	47	SER	HN	8387.0
HB2	46	TYR	HD1	13717.2
HB3	46	TYR	HD1	17581.2
HD1	46	TYR	HE1	66603.6
HD1	46	TYR	HE2	25460.4
HE1	46	TYR	HD2	35649.6
HE2	46	TYR	HD2	143724.0
HE2	46	TYR	HH	6149.0
HE2	48	TYR	HA	21184.8
HD2	48	TYR	HA	19891.2
HH	79	LYS+	HE3	1959.0
47	SER			
HN	47	SER	HA	4633.0
HN	47	SER	HB2	5690.0
HN	47	SER	HB3	9881.0
HA	47	SER	HB2	32113.2
HA	47	SER	HB3	76759.2
HA	48	TYR	HN	27590.0
HB2	48	TYR	HN	1604.0
48	TYR			
HN	48	TYR	HA	826.0
HN	48	TYR	HB2	17060.0
HN	48	TYR	HB3	11290.0
HA	48	TYR	HB2	18732.0
HA	48	TYR	HB3	7161.0
HA	49	THR	HN	8034.0
HB2	48	TYR	HD1	12096.0
HB2	48	TYR	HE2	5538.1
HB2	48	TYR	HD2	15338.4
HB2	49	THR	HN	5008.0
HB2	53	ILE	HN	1205.0
HB3	48	TYR	HD1	5437.3
HB3	48	TYR	HE2	7367.6
HB3	48	TYR	HD2	11440.8
HB3	49	THR	HN	7955.0
HD1	48	TYR	HE1	28996.8
HD1	48	TYR	HH	2063.1
HE1	48	TYR	HD2	39345.6
HE1	48	TYR	HH	13225.0
HE2	48	TYR	HH	8450.2
HE2	49	THR	HN	3196.0
49	THR			
HN	49	THR	HA	3172.0
HN	49	THR	HB	3766.0
HN	49	THR	QG2	23310.0
HN	49	THR	HG1	1944.0
HN	52	ASN	HN	2779.0
HN	52	ASN	HB2	6772.0
HN	52	ASN	HB3	6097.0
HA	49	THR	HB	41697.6
HA	49	THR	QG2	83143.2
HA	50	ASP-	HN	47357.0
HB	49	THR	QG2	74944.8
HB	50	ASP-	HN	17859.5
HB	51	ALA	HN	21965.0
HB	51	ALA	QB	4331.9
QG2	50	ASP-	HN	4301.0
QG2	78	THR	HA	19513.2
HG1	52	ASN	HN	14320.0
50	ASP-			
HN	50	ASP-	HA	10562.8
HN	50	ASP-	HB2	23430.0
HN	50	ASP-	HB3	5590.0
HN	51	ALA	HN	11250.0
HA	50	ASP-	HB2	7808.0
HA	50	ASP-	HB3	17136.0
HA	51	ALA	HN	5441.8
HA	53	ILE	HN	2583.0
HA	53	ILE	HB	20034.0
HA	53	ILE	QD1	25410.0
HB2	51	ALA	HN	5905.0
HB3	51	ALA	HN	9990.0
HB3	51	ALA	HA	2043.0
51	ALA			
HN	51	ALA	HA	18160.0
HN	51	ALA	QB	25840.0
HN	52	ASN	HN	12270.0
HA	51	ALA	QB	95960.0
HA	52	ASN	HN	3753.0
HA	54	LYS+	HN	771.1
HA	54	LYS+	HG3	8777.0

HA	75	ILE	QG2	6274.8
QB	52	ASN	HN	22100.0
QB	77	GLY	HA2	21285.6
QB	78	THR	HA	760.4
QB	78	THR	HB	3516.2
52 ASN				
HN	52	ASN	HA	5038.0
HN	52	ASN	HB2	17430.0
HN	52	ASN	HB3	24010.0
HN	53	ILE	HN	12750.0
HN	75	ILE	QG2	2082.0
HA	52	ASN	HB2	38976.0
HA	52	ASN	HB3	10878.0
HA	52	ASN	HD21	4819.0
HA	55	LYS+	HN	8948.0
HA	55	LYS+	HB2	5637.0
HA	55	LYS+	HB3	16090.0
HA	75	ILE	QG2	51156.0
HB2	52	ASN	HD21	16810.0
HB2	52	ASN	HD22	4260.0
HB2	53	ILE	HN	3400.0
HB2	75	ILE	QG2	29601.6
HB3	52	ASN	HD21	9878.0
HB3	52	ASN	HD22	3095.0
HB3	53	ILE	HN	13490.0
HB3	75	ILE	QG2	12087.6
HD21	59	TRP	HZ2	7658.0
HD21	75	ILE	QG2	1957.0
HD22	59	TRP	HE1	1049.0
HD22	59	TRP	HZ2	20450.0
53 ILE				
HN	53	ILE	HA	14220.0
HN	53	ILE	HB	34680.0
HN	53	ILE	QG2	11700.0
HN	53	ILE	HG12	11040.0
HN	53	ILE	HG13	25280.0
HN	53	ILE	QD1	5247.0
HN	54	LYS+	HN	13310.0
HN	54	LYS+	HG2	1775.0
HA	53	ILE	HB	12340.0
HA	53	ILE	QG2	42260.0
HA	53	ILE	HG13	7545.0
HA	53	ILE	QD1	10466.4
HA	54	LYS+	HN	3232.0
HA	55	LYS+	HN	4478.0
HB	53	ILE	QG2	101400.0
HB	53	ILE	HG12	27232.8
HB	53	ILE	QD1	62143.2
HB	54	LYS+	HN	29050.0
HB	54	LYS+	HG2	9304.0
QG2	54	LYS+	HN	13730.0
QG2	54	LYS+	HA	2880.0
QG2	54	LYS+	HG2	30290.0
HG12	53	ILE	QD1	428652.0
HG13	53	ILE	QD1	62974.8
54 LYS+				
HN	54	LYS+	HA	20390.0
HN	54	LYS+	HB2	45560.0
HN	54	LYS+	HB3	3787.0
HN	54	LYS+	HG2	26900.0
HN	54	LYS+	HG3	11280.0
HN	55	LYS+	HN	11370.0
HN	56	ASN	HN	1889.0
HA	54	LYS+	HB2	38910.0
HA	54	LYS+	HG2	21450.0
HA	54	LYS+	HG3	18270.0
HA	55	LYS+	HN	5291.0
HB2	55	LYS+	HN	8222.0
HB3	55	LYS+	HN	1304.0
55 LYS+				
HN	55	LYS+	HA	14990.0
HN	55	LYS+	HB2	14640.0
HN	55	LYS+	HB3	38790.0
HN	55	LYS+	HG2	1313.0
HN	56	ASN	HN	30980.0
HA	55	LYS+	HB2	40470.0
HA	55	LYS+	HB3	12350.0
HA	55	LYS+	HG2	1625.0
HA	56	ASN	HN	4578.0
HB2	55	LYS+	HD2	14054.4
HB2	56	ASN	HN	1286.0
HB2	75	ILE	QG2	31380.0
HB3	56	ASN	HN	2604.0
HB3	57	VAL	HN	3113.0
HB3	57	VAL	QG2	5571.0
HB3	75	ILE	QG2	15440.0
HG2	57	VAL	QG2	31980.0
HG2	75	ILE	QG2	37260.0
HG2	75	ILE	QD1	43580.0
HD2	57	VAL	QG2	8923.0
56 ASN				
HN	56	ASN	HA	34890.0
HN	56	ASN	HB2	18800.0
HN	56	ASN	HB3	4954.0
HA	56	ASN	HB2	25569.6
HA	56	ASN	HB3	11928.0
HA	57	VAL	HN	24390.0
HB2	56	ASN	HD22	94.2
HB3	56	ASN	HD22	2689.0
HN	57	VAL	HA	13540.0
HN	57	VAL	HB	13070.0
HN	57	VAL	QG1	10310.0
HN	57	VAL	QG2	20360.0
HA	57	VAL	HB	10710.0
HA	57	VAL	QG1	16480.8
HA	57	VAL	QG2	20460.0
HA	58	LEU	HN	70430.0
HB	57	VAL	QG1	107000.0
HB	57	VAL	QG2	41610.0
HB	58	LEU	HN	1157.0
HB	59	TRP	HZ2	2442.7
QG1	58	LEU	HN	8919.0
QG1	59	TRP	HA	3522.1
QG1	59	TRP	HE3	8786.0
QG1	59	TRP	HZ3	6828.4
QG1	59	TRP	HZ2	5607.8
QG1	63	ASN	HD21	8463.0
QG1	63	ASN	HD22	3468.0
QG1	74	TYR	CZ	9700.0*2
QG2	58	LEU	HN	12610.0
QG2	59	TRP	HE3	3570.0
QG2	59	TRP	HZ3	3090.0
QG2	59	TRP	HZ2	3103.0
QG2	59	TRP	HH2	2504.0
QG2	63	ASN	HD21	9986.0
QG2	63	ASN	HD22	7602.0
QG2	74	TYR	CG	1763.5*2
QG2	74	TYR	CZ	13350.0*2
QG2	75	ILE	QD1	8180.8
58 LEU				
HN	58	LEU	HA	6048.0
HN	58	LEU	HB2	18590.0
HN	58	LEU	HB3	15950.0
HN	58	LEU	QD1	1025.0
HN	58	LEU	QD2	1754.0

HA	58	LEU	HB2	30580.0
HA	58	LEU	HB3	8916.0
HA	58	LEU	QD1	30256.8
HA	58	LEU	QD2	6851.0
HB2	58	LEU	QD1	26460.0
HB2	58	LEU	QD2	22520.4
HB3	58	LEU	QD1	24850.0
HB3	58	LEU	QD2	58371.6
QD1	60	ASP-	HB3	11530.0
QD2	60	ASP-	HB3	3089.5
59	TRP			
HA	59	TRP	HB2	16153.2
HA	59	TRP	HB3	29509.2
HA	59	TRP	HE3	5671.7
HB2	59	TRP	HD1	9609.6
HB2	59	TRP	HE3	6957.0
HB2	60	ASP-	HN	516.0
HB3	59	TRP	HD1	16245.6
HB3	59	TRP	HE3	23503.2
HB3	60	ASP-	HN	692.0
HD1	59	TRP	HE1	27460.0
HD1	64	MET	QE	1210.0
HE3	59	TRP	HZ3	38690.4
HE3	59	TRP	HH2	9189.6
HE3	60	ASP-	HN	5034.0
HE3	64	MET	HN	9526.0
HE3	64	MET	HA	3359.0
HE3	64	MET	HB2	7027.0
HE3	64	MET	HB3	2955.0
HE3	64	MET	QE	11163.6
HE3	67	TYR	HB2	3040.0
HE1	59	TRP	HZ2	17550.0
HZ3	59	TRP	HZ2	9004.8
HZ3	59	TRP	HH2	63058.8
HZ3	63	ASN	HB3	2023.0
HZ3	64	MET	HN	999.0
HZ3	64	MET	HA	25074.0
HZ3	64	MET	QE	1421.0
HZ3	67	TYR	HN	11220.0
HZ3	67	TYR	HA	4086.6
HZ3	67	TYR	HB2	31332.0
HZ3	67	TYR	HB3	28030.8
HZ3	74	TYR	CG	1228.0*2
HZ2	59	TRP	HH2	59094.0
HZ2	75	ILE	QD1	11088.0
HH2	64	MET	QE	1232.0
HH2	67	TYR	HB2	16330.0
HH2	67	TYR	HB3	14011.2
HH2	74	TYR	CG	13695.0*2
HH2	74	TYR	CZ	838.5*2
HH2	75	ILE	HG12	8986.0
HH2	75	ILE	QD1	23167.2
60	ASP-			
HN	63	ASN	HB3	1043.0
HN	63	ASN	HD21	5165.0
HN	63	ASN	HD22	2933.0
HN	64	MET	HN	722.0
HN	64	MET	HB2	860.0
HA	60	ASP-	HB2	33512.4
HA	60	ASP-	HB3	54036.0
HB2	61	GLU-	HN	18600.0
HB2	62	ASN	HN	6771.0
HB3	61	GLU-	HN	2018.0
61	GLU-			
HN	61	GLU-	HA	13280.0
HN	61	GLU-	HB3	31810.0
HN	61	GLU-	HG2	13280.0
HN	61	GLU-	HG3	11330.0
HN	62	ASN	HN	10860.0
HA	61	GLU-	HB3	30140.0
HA	61	GLU-	HG2	33110.0
HA	61	GLU-	HG3	20420.4
HA	62	ASN	HN	5114.0
HA	64	MET	HN	3707.0
HA	95	ILE	QG2	19714.8
HA	95	ILE	HG12	5845.6
HA	95	ILE	QD1	11860.8
HB2	62	ASN	HN	16260.0
HB2	95	ILE	QG2	9408.0
HB3	61	GLU-	HG2	136700.0
HB3	61	GLU-	HG3	119700.0
HG2	62	ASN	HN	3596.0
HG3	95	ILE	QG2	5733.0
62	ASN			
HN	62	ASN	HB2	20650.0
HN	62	ASN	HB3	11800.0
HN	62	ASN	HD21	4296.0
HN	63	ASN	HN	17300.0
HN	64	MET	HN	2402.0
HA	62	ASN	HB2	6363.8
HA	62	ASN	HB3	17018.4
HB2	62	ASN	HD21	11240.0
HB2	62	ASN	HD22	16110.0
HB3	62	ASN	HD21	8390.0
HB3	62	ASN	HD22	12050.0
HB3	63	ASN	HN	10260.0
63	ASN			
HN	63	ASN	HA	2527.0
HN	63	ASN	HB2	10080.0
HN	63	ASN	HB3	5017.0
HN	64	MET	HN	23330.0
HA	63	ASN	HB2	15355.2
HA	63	ASN	HB3	15372.0
HA	66	GLU-	HB2	8392.4
HA	66	GLU-	HB3	10374.0
HB2	63	ASN	HD21	20920.0
HB2	63	ASN	HD22	26090.0
HB2	64	MET	HN	2826.0
HB3	63	ASN	HD21	6542.0
HB3	63	ASN	HD22	4506.0
HB3	64	MET	HN	2795.0
64	MET			
HN	64	MET	HA	6771.0
HN	64	MET	HB2	45250.0
HN	64	MET	HB3	24490.0
HN	64	MET	QE	5094.0
HN	65	SER	HN	12060.0
HA	67	TYR	HN	5569.0
HA	67	TYR	HB2	11566.8
HB2	64	MET	QE	29360.0
QE	98	LEU	QD1	32516.4
65	SER			
HN	65	SER	HA	7253.0
HN	65	SER	HB2	8959.0
HN	65	SER	HB3	10190.0
HN	66	GLU-	HN	14930.0
HN	67	TYR	HN	832.9
HN	95	ILE	QD1	15770.0
HA	65	SER	HB2	31592.4
HA	65	SER	HB3	7693.6
HA	68	LEU	HN	1185.0
HA	68	LEU	QD2	5024.9
HA	95	ILE	QD1	24788.4
HB2	95	ILE	QD1	37111.2
66	GLU-			
HN	66	GLU-	HA	10110.0

HN	66	GLU-	HB2	18260.0
HN	66	GLU-	HB3	20330.0
HN	66	GLU-	HG2	2739.0
HN	67	TYR	HN	13940.0
HN	68	LEU	HN	2011.0
HA	66	GLU-	HB2	22772.4
HA	66	GLU-	HB3	14649.6
HA	66	GLU-	HG2	13599.6
HA	66	GLU-	HG3	18992.4
HA	67	TYR	HN	1012.0
HB2	66	GLU-	HG2	35448.0
HB2	67	TYR	HN	5423.0
HB2	74	TYR	CG	2335.0*2
HB2	74	TYR	CZ	3817.0*2
HB3	66	GLU-	HG2	42218.4
HB3	67	TYR	HN	10710.0
HB3	74	TYR	CG	2690.0*2
HB3	74	TYR	CZ	3719.0*2
HG2	74	TYR	CG	7200.0*2
HG2	74	TYR	CZ	4887.0*2
HG3	67	TYR	HN	7516.0
HG3	74	TYR	CG	1609.4*2
67 TYR				
HN	67	TYR	HA	3171.0
HN	67	TYR	HB2	25110.0
HN	67	TYR	HB3	31140.0
HN	74	TYR	CZ	482.6*2
HA	67	TYR	HB2	11062.8
HA	67	TYR	HB3	9693.6
HA	71	PRO	HA	8559.6
HA	74	TYR	CG	5607.0*2
HB2	68	LEU	HN	3528.0
HB2	74	TYR	CG	8260.0*2
HB2	74	TYR	CZ	3063.5*2
HB3	68	LEU	HN	11330.0
HB3	74	TYR	CG	3994.2*2
HD1	67	TYR	HE2	3249.0
HD1	74	TYR	HB3	1153.0
68 LEU				
HN	68	LEU	HA	858.0
HN	68	LEU	HB2	21920.0
HN	68	LEU	HB3	4944.0
HN	68	LEU	HG	9485.0
HN	68	LEU	QD1	2893.0
HN	68	LEU	QD2	15540.0
HN	69	THR	HN	9319.0
HA	68	LEU	HB2	22537.2
HA	68	LEU	HB3	15792.0
HA	68	LEU	HG	13683.6
HA	68	LEU	QD1	44058.0
HA	68	LEU	QD2	7337.4
HA	71	PRO	HG2	298.1
HA	71	PRO	HG3	8526.0
HA	71	PRO	HD3	14473.2
HA	82	PHE	QR	1504.6*5
HB2	68	LEU	QD1	8066.0
HB2	68	LEU	QD2	19970.0
HB2	69	THR	HN	917.0
HB3	68	LEU	QD1	16700.0
HB3	68	LEU	QD2	38665.2
HB3	69	THR	HN	8351.0
HB3	82	PHE	QR	295.4*5
HB3	85	LEU	HN	3482.0
HB3	85	LEU	QD1	2059.7
HG	68	LEU	QD1	27340.0
HG	68	LEU	QD2	45950.0
QD1	71	PRO	HG2	3002.2
QD1	71	PRO	HG3	4216.8
QD1	82	PHE	HZ	775.7
QD1	82	PHE	QR	4254.0*5
QD1	85	LEU	QD1	2572.0
QD2	82	PHE	QR	6614.0*5
QD2	85	LEU	QD1	11617.2
QD2	85	LEU	QD2	5251.0
QD2	91	ARG+	HA	10021.2
QD2	94	LEU	HB3	6996.0
QD2	94	LEU	HG	23444.4
QD2	94	LEU	QD1	20770.0
QD2	95	ILE	HA	864.4
QD2	95	ILE	HG12	2687.2
QD2	95	ILE	QD1	23030.0
69 THR				
HN	69	THR	HA	2670.0
HN	70	ASN	HN	20420.0
HA	69	THR	HB	11365.2
HA	69	THR	QG2	40412.4
HA	84	GLY	HA2	9735.6
HA	85	LEU	HN	3643.0
HB	69	THR	QG2	14742.0
QG2	86	LYS+	HA	21747.6
70 ASN				
HN	70	ASN	HA	4210.0
HN	70	ASN	HB2	6146.0
HN	71	PRO	HA	1673.0
HN	71	PRO	HD2	2523.0
HN	71	PRO	HD3	5713.0
HA	70	ASN	HB2	24259.2
HA	70	ASN	HB3	27426.0
HA	71	PRO	HG2	2374.0
HA	71	PRO	HG3	1474.2
HA	71	PRO	HD2	38379.6
HA	71	PRO	HD3	93390.0
HA	72	TML	HN	1533.0
HA	84	GLY	HA1	12516.0
HA	84	GLY	HA2	19395.6
HB2	73	LYS+	HN	1473.0
HB2	73	LYS+	HB2	6103.4
HB3	73	LYS+	HB3	15262.8
71 PRO				
HA	71	PRO	HB2	41748.0
HA	71	PRO	HB3	33200.0
HA	71	PRO	HG2	9480.0
HA	71	PRO	HG3	22495.2
HA	71	PRO	HD2	5811.1
HA	71	PRO	HD3	8450.4
HA	73	LYS+	HN	5510.0
HA	74	TYR	HN	2076.0
HA	74	TYR	HB2	7034.2
HA	74	TYR	HB3	8618.4
HA	75	ILE	HG12	10382.4
HA	78	THR	QG2	4137.0
HB2	71	PRO	HG2	16531.2
HB2	71	PRO	HG3	20563.2
HB2	71	PRO	HD2	34532.4
HB2	71	PRO	HD3	2756.9
HB2	72	TML	HN	20620.0
HB2	78	THR	QG2	31206.0
HB3	71	PRO	HG2	17380.0
HB3	71	PRO	HG3	18540.0
HB3	71	PRO	HD2	7137.5
HB3	71	PRO	HD3	5453.3
HB3	72	TML	HN	3206.0
HB3	78	THR	QG2	8912.4
HG2	71	PRO	HD2	16909.2
HG2	71	PRO	HD3	17505.6
HG2	82	PHE	QR	1973.6*5

HG3	71	PRO	HD2	20252.4
HG3	71	PRO	HD3	20874.0
HD2	72	TML	HN	5365.0
HD2	84	GLY	HA1	10852.8
HD2	84	GLY	HA2	10189.2
HD3	72	TML	HN	579.0
HD3	82	PHE	QR	764.4*5
HD3	84	GLY	HA1	22722.0
HD3	84	GLY	HA2	25124.4
72	TML			
HN	72	TML	HA	836.0
HN	72	TML	HB2	7012.0
HN	72	TML	HB3	5867.0
HN	73	LYS+	HN	13320.0
HN	74	TYR	HN	4520.0
HN	78	THR	QG2	10270.0
HA	72	TML	HB2	11870.0
HA	72	TML	HB3	4946.8
HA	73	LYS+	HN	589.0
HA	78	THR	HN	106.0
HB2	72	TML	HE2	6304.0
HB3	72	TML	QQH	21140.0
HG2	72	TML	HE2	11360.0
HG2	72	TML	QQH	21820.0
HG2	73	LYS+	HN	9328.0
HG3	72	TML	HE2	9540.0
HE2	72	TML	QQH	38750.0
QQH	81	ALA	HA	17230.0
QQH	81	ALA	QB	10140.0
QQH	82	PHE	HN	5005.0
73	LYS+			
HN	73	LYS+	HB2	31710.0
HN	73	LYS+	HB3	26370.0
HN	73	LYS+	HG2	5077.0
HN	73	LYS+	HG3	2888.0
HN	73	LYS+	HD2	5259.0
HN	74	TYR	HN	49880.0
HN	75	ILE	HN	13110.0
HA	73	LYS+	HB2	38396.4
HA	73	LYS+	HB3	13246.8
HA	73	LYS+	HG2	16228.8
HA	73	LYS+	HG3	20017.2
HA	73	LYS+	HD2	7376.0
HA	73	LYS+	HE2	3612.0
HB2	73	LYS+	HE2	21848.4
HB2	74	TYR	HN	5219.0
HB3	73	LYS+	HE2	880.3
HB3	74	TYR	HN	13220.0
HG2	73	LYS+	HE2	9547.0
HD2	73	LYS+	HE2	57615.6
74	TYR			
HN	75	ILE	HN	13170.0
HA	74	TYR	HB2	21000.0
HA	74	TYR	HB3	11844.0
HA	74	TYR	CG	4321.8*2
HB2	74	TYR	CG	14460.6*2
HB3	74	TYR	CG	13494.6*2
HB3	75	ILE	HG12	18211.2
CG	74	TYR	CZ	24985.0*2*2
CG	75	ILE	HN	2551.5*2
CG	75	ILE	HB	3631.0*2
CG	75	ILE	QG2	1148.3*2
CG	75	ILE	HG13	17304.0*2
CG	75	ILE	QD1	13108.2*2
CZ	75	ILE	HG13	682.5*2
CZ	75	ILE	QD1	2047.9*2
75	ILE			
HN	75	ILE	HB	14770.0
HN	75	ILE	QG2	6691.0
HN	75	ILE	HG12	4216.0
HN	75	ILE	HG13	11310.0
HN	75	ILE	QD1	6055.0
HN	78	THR	QG2	13240.0
HA	75	ILE	QG2	12398.4
HA	75	ILE	HG12	11583.6
HA	75	ILE	QD1	5547.4
HB	78	THR	HB	40084.8
QG2	75	ILE	QD1	60940.0
QG2	78	THR	HA	7623.0
QG2	78	THR	HB	31441.2
76	PRO			
HA	76	PRO	HB2	18463.2
HA	76	PRO	HB3	107520.0
HA	76	PRO	HG2	17623.2
HA	76	PRO	HD2	5096.3
HA	77	GLY	HN	2668.0
HB3	76	PRO	HG2	330792.0
HB3	76	PRO	HD2	62739.6
HG2	76	PRO	HD2	82454.4
77	GLY			
HN	77	GLY	HA1	993.0
HN	78	THR	HN	321.0
HN	78	THR	QG2	16590.0
78	THR			
HN	78	THR	HB	5683.0
HN	78	THR	QG2	70850.0
HA	78	THR	HB	29576.4
HA	78	THR	QG2	17833.2
HB	78	THR	QG2	297864.0
QG2	79	LYS+	HN	7656.0
QG2	80	ALA	HN	5867.0
79	LYS+			
HN	79	LYS+	HG3	6735.0
HN	79	LYS+	HD2	2485.0
HN	79	LYS+	HE3	2193.0
HN	80	ALA	HN	10250.0
HB2	79	LYS+	HG2	38700.0
HB2	79	LYS+	HG3	17650.0
HB2	79	LYS+	HE3	3087.0
HB3	79	LYS+	HG2	10140.0
HB3	79	LYS+	HG3	6171.0
HB3	79	LYS+	HE3	4340.0
HG2	79	LYS+	HE3	17830.0
HG2	80	ALA	HN	5643.0
HG3	79	LYS+	HD2	16720.0
HG3	79	LYS+	HE3	38350.0
HG3	80	ALA	HN	14060.0
HG3	80	ALA	HA	1365.0
HD2	79	LYS+	HE3	50490.0
80	ALA			
HN	80	ALA	HA	12840.0
HN	80	ALA	QB	23040.0
HA	80	ALA	QB	11830.0
HA	81	ALA	HN	52050.0
HA	81	ALA	HA	386.0
HA	82	PHE	HN	964.0
81	ALA			
HN	81	ALA	HA	12160.0
HN	81	ALA	QB	33420.0
HN	82	PHE	HN	670.0
HA	81	ALA	QB	90828.0
HA	82	PHE	HN	15090.0
82	PHE			
HN	82	PHE	HB2	5356.0
HZ	82	PHE	QR	24860.0*5
HZ	84	GLY	HA1	4652.0

```
HZ      85 LEU  HN      2952.0
HZ      85 LEU  HA      3272.6
HZ      85 LEU  HB3    37580.0
HZ      85 LEU  HG     10130.4
HZ      85 LEU  QD1    12900.0
HZ      85 LEU  QD2    26120.0
QR      84 GLY  HA1     3922.8*5
QR      84 GLY  HA2     1953.8*5
QR      85 LEU  HN      2298.0*5
QR      85 LEU  HA      1137.7*5
QR      85 LEU  HB3     3596.9*5
QR      85 LEU  HG      4516.0*5
QR      85 LEU  QD1     4538.0*5
QR      85 LEU  QD2     1530.0*5
 83 GLY
HN      83 GLY  HA2     5407.3
HA2     85 LEU  HN       955.0
 84 GLY
HA1     85 LEU  HN      6531.0
HA2     85 LEU  HN      8510.0
 85 LEU
HN      85 LEU  HB2    10890.0
HN      85 LEU  HB3     4044.0
HN      85 LEU  HG     17150.0
HN      85 LEU  QD1     4507.0
HN      85 LEU  QD2     2021.0
HA      85 LEU  HB2    22369.2
HA      85 LEU  HB3    16514.4
HA      85 LEU  HG     27837.6
HA      85 LEU  QD1    10466.4
HA      85 LEU  QD2    41882.4
HB2     85 LEU  QD1     9752.4
HB3     85 LEU  QD1    27115.2
HB3     85 LEU  QD2    41420.4
HG      85 LEU  QD1    65226.0
QD1     91 ARG+ HA     26000.0
QD1     94 LEU  HG     21378.0
QD2     94 LEU  QD2     2210.9
 86 LYS+
HN      86 LYS+ HA       376.0
HN      86 LYS+ HB2     6640.0
HN      86 LYS+ HB3    10320.0
HN      86 LYS+ HG2    13910.0
HN      87 LYS+ HN     20940.0
HA      86 LYS+ HB2     6252.1
HA      86 LYS+ HB3    27711.6
HA      86 LYS+ HG2    17371.2
HA      86 LYS+ HG3    31021.2
HA      86 LYS+ HD2    12087.6
HA      86 LYS+ HD3    13246.8
HA      86 LYS+ HE2     6909.0
HB2     87 LYS+ HN      5364.0
HB3     87 LYS+ HN      5807.0
 87 LYS+
HN      87 LYS+ HA      2873.0
HN      87 LYS+ HB2     6907.0
HN      87 LYS+ HB3    25880.0
HN      87 LYS+ HG2    13260.0
HN      87 LYS+ HG3    10600.0
HA      87 LYS+ HB2    43528.8
HA      87 LYS+ HB3    16816.8
HB2     90 ASP- HN      5394.0
HB3     89 LYS+ HN     33370.0
HB3     90 ASP- HN     24440.0
 88 GLU-
HN      89 LYS+ HN     12840.0
 89 LYS+
HN      89 LYS+ HA     13580.0
HN      89 LYS+ HB2    34610.0
HN      89 LYS+ HG2    14300.0
HN      89 LYS+ HD3    24960.0
HN      90 ASP- HN     12330.0
HA      89 LYS+ HB2    30660.0
HA      89 LYS+ HB3    18660.0
HA      89 LYS+ HG2    18958.8
HA      89 LYS+ HG3    26073.6
HA      89 LYS+ HD2    15808.8
HA      89 LYS+ HD3    27980.4
HA      90 ASP- HN      4185.0
HA      92 ASN  HN      4722.0
HA      93 ASP- HN      2181.0
HB2     90 ASP- HN     17430.0
HB3     90 ASP- HN     12360.0
HG3     90 ASP- HN      8872.0
 90 ASP-
HN      90 ASP- HA     17730.0
HN      90 ASP- HB2    15290.0
HN      90 ASP- HB3     6503.0
HN      91 ARG+ HN     21430.0
HN      92 ASN  HN      1421.0
HA      90 ASP- HB2    15859.2
HA      90 ASP- HB3    19420.8
HA      91 ARG+ HN      3970.0
HA      93 ASP- HN     12020.0
HB2     91 ARG+ HN      6921.0
HB3     91 ARG+ HN     11950.0
 91 ARG+
HN      91 ARG+ HA     11100.0
HN      91 ARG+ HB2    19960.0
HN      91 ARG+ HB3    19850.0
HN      91 ARG+ HG2     9066.0
HN      92 ASN  HN      9949.0
HA      91 ARG+ HB2    21361.2
HA      91 ARG+ HB3    13880.0
HA      91 ARG+ HG2    58086.0
HA      91 ARG+ HD2     3473.4
HA      92 ASN  HN      4820.0
HA      94 LEU  HN      3891.0
HA      94 LEU  HB2     8408.4
HA      94 LEU  HB3    46317.6
HA      95 ILE  QD1     9517.2
HB2     91 ARG+ HG2    84840.0
HB2     92 ASN  HN      4686.0
HB2     95 ILE  HG12    9307.2
HB3     91 ARG+ HG2    27274.8
HB3     92 ASN  HN     12690.0
HB3     95 ILE  QD1    13582.8
 92 ASN
HN      92 ASN  HA      1609.0
HN      92 ASN  HB3    17780.0
HN      94 LEU  HB3     2868.0
HA      92 ASN  HB3    20302.8
HA      93 ASP- HN      1555.0
HA      95 ILE  QD1     9433.2
HB3     93 ASP- HN     16720.0
 93 ASP-
HN      93 ASP- HA      7747.0
HN      93 ASP- HB2    23990.0
HN      93 ASP- HB3     5800.0
HN      94 LEU  HN     19390.0
HA      93 ASP- HB2    17211.6
HA      93 ASP- HB3    10060.0
HA      94 LEU  HN      4062.0
HB2     94 LEU  HN      5835.0
HB3     94 LEU  HN     10250.0
 94 LEU
```

HN	94	LEU	HA	9368.0
HN	94	LEU	HB3	34510.0
HN	94	LEU	HG	5513.0
HN	94	LEU	QD1	4805.0
HN	94	LEU	QD2	6047.0
HN	95	ILE	HN	14000.0
HA	94	LEU	HB2	8412.0
HA	94	LEU	HB3	13120.8
HA	94	LEU	HG	12154.8
HA	94	LEU	QD1	6930.8
HA	94	LEU	QD2	25980.0
HA	95	ILE	HN	4580.0
HA	97	TYR	HN	2713.0
HB2	94	LEU	HG	127700.0
HB2	94	LEU	QD1	23116.8
HB2	94	LEU	QD2	12776.4
HB2	95	ILE	HN	30780.0
HB3	94	LEU	HG	49098.0
HB3	94	LEU	QD1	32785.2
HB3	94	LEU	QD2	30181.2
HB3	95	ILE	HN	2349.0
HB3	95	ILE	HG12	8828.4
HG	94	LEU	QD1	36346.8
HG	94	LEU	QD2	28719.6
HG	95	ILE	HN	15900.0
HG	95	ILE	HA	7594.4
QD1	97	TYR	HN	835.0
QD1	98	LEU	HN	9146.0
QD1	98	LEU	HG	13188.0
QD1	98	LEU	QD1	12079.2
QD2	98	LEU	QD1	3079.4
95	ILE			
HN	95	ILE	HA	14550.0
HN	95	ILE	HB	37930.0
HN	95	ILE	QG2	11160.0
HN	95	ILE	HG12	18640.0
HN	95	ILE	QD1	17680.0
HN	96	THR	HN	8409.0
HA	95	ILE	HB	11986.8
HA	95	ILE	QG2	24864.0
HA	95	ILE	HG12	12734.4
HA	95	ILE	QD1	4210.0
HA	96	THR	HN	3862.0
HA	98	LEU	HN	12300.0
HA	98	LEU	HB2	16880.0
HA	98	LEU	HB3	16060.0
HA	98	LEU	QD1	1963.1
HA	98	LEU	QD2	1945.4
HA	99	LYS+	HN	2217.0
HB	95	ILE	QG2	37161.6
HB	95	ILE	HG12	36900.0
HB	95	ILE	QD1	50005.2
HB	96	THR	HN	20060.0
QG2	95	ILE	QD1	53146.8
QG2	96	THR	HN	11840.0
QG2	96	THR	HA	1305.4
QG2	98	LEU	HB2	27070.0
QG2	99	LYS+	HN	1928.0
QG2	99	LYS+	HB2	30928.8
QG2	99	LYS+	HB3	7320.6
HG12	95	ILE	QD1	22965.6
HG12	98	LEU	QD2	8946.0
QD1	96	THR	HN	2372.0
96	THR			
HN	96	THR	HA	8831.0
HN	97	TYR	HN	14760.0
HA	96	THR	HB	12776.4
HA	96	THR	QG2	33310.0
HA	97	TYR	HN	3204.0
HA	99	LYS+	HN	1087.0
HA	99	LYS+	HB2	6036.2
HA	99	LYS+	HB3	31458.0
HB	96	THR	QG2	75616.8
QG2	97	TYR	HN	7808.0
QG2	97	TYR	HA	18916.8
97	TYR			
HN	97	TYR	HA	20590.0
HN	97	TYR	HB2	16630.0
HN	97	TYR	HB3	14840.0
HN	98	LEU	HN	24480.0
HA	97	TYR	HB2	53894.4
HA	97	TYR	HB3	13381.2
HA	98	LEU	HN	5489.0
HA	100	LYS+	HN	4587.0
HA	101	ALA	HN	2678.0
HB2	98	LEU	HN	2754.0
HB3	98	LEU	HN	20340.0
HD1	97	TYR	HE1	1700.0
HD1	97	TYR	HE2	9697.0
HE1	97	TYR	HD2	3508.0
HE2	97	TYR	HD2	8246.3
98	LEU			
HN	98	LEU	HA	11220.0
HN	98	LEU	HB2	34480.0
HN	98	LEU	HB3	12120.0
HN	98	LEU	HG	4398.0
HN	98	LEU	QD1	4309.0
HN	98	LEU	QD2	572.0
HN	99	LYS+	HN	21910.0
HN	101	ALA	QB	1967.0
HA	98	LEU	HB2	19303.2
HA	98	LEU	HB3	4032.8
HA	98	LEU	QD1	5180.3
HA	98	LEU	QD2	24343.2
HA	99	LYS+	HN	3723.0
HA	100	LYS+	HN	3548.0
HA	101	ALA	HN	6848.0
HA	101	ALA	QB	26830.0
HA	102	SER	HN	3996.0
HB2	98	LEU	QD1	74410.0
HB2	98	LEU	QD2	61311.6
HB2	99	LYS+	HN	14390.0
HB3	98	LEU	QD1	37210.0
HB3	98	LEU	QD2	32432.4
HB3	99	LYS+	HN	19630.0
HG	98	LEU	QD1	16580.0
HG	98	LEU	QD2	20714.4
QD1	99	LYS+	HN	1599.0
99	LYS+			
HN	99	LYS+	HA	17400.0
HN	99	LYS+	HB2	22070.0
HN	99	LYS+	HB3	47840.0
HN	99	LYS+	HG2	7700.0
HN	99	LYS+	HG3	10920.0
HN	100	LYS+	HN	15860.0
HA	99	LYS+	HB2	33465.6
HA	99	LYS+	HB3	11104.8
HA	99	LYS+	HG2	21450.0
HA	99	LYS+	HG3	18340.0
HA	99	LYS+	HD2	13040.0
HA	100	LYS+	HN	4627.0
HB2	99	LYS+	HG2	20538.0
HB2	99	LYS+	HG3	16926.0
HB2	99	LYS+	HD2	14977.2
HB2	99	LYS+	HD3	45301.2
HB3	99	LYS+	HG2	11650.8

HB3	99	LYS+	HG3	27140.4
HB3	99	LYS+	HD2	16245.6
HB3	99	LYS+	HD3	21546.0
HB3	100	LYS+	HN	18840.0
HG2	99	LYS+	HD2	10718.4
HG2	99	LYS+	HD3	11010.0
HG2	100	LYS+	HG2	10978.8
HG3	99	LYS+	HD3	30860.0
HG3	100	LYS+	HN	9618.0
100 LYS+				
HN	100	LYS+	HA	18140.0
HN	100	LYS+	HB2	29630.0
HN	100	LYS+	HB3	18660.0
HN	100	LYS+	HG2	34920.0
HN	100	LYS+	HD2	4260.0
HN	101	ALA	HN	38170.0
HN	101	ALA	QB	6212.0
HN	102	SER	HN	4045.0
HA	100	LYS+	HB2	13790.0
HA	100	LYS+	HB3	18910.0
HA	100	LYS+	HG2	29500.8
HA	101	ALA	HN	3551.0
HA	103	GLU-	HB2	1634.0
HB2	100	LYS+	HG2	170604.0
HB2	101	ALA	HN	8669.0
HB3	101	ALA	HN	9215.0
HG2	100	LYS+	HD2	109284.0
HG2	100	LYS+	HE2	46779.6
HG2	101	ALA	HN	4147.0
HD2	100	LYS+	HE2	20412.0
101 ALA				
HN	101	ALA	HA	12190.0
HN	101	ALA	QB	67320.0
HN	102	SER	HN	19940.0
HN	103	GLU-	HN	1856.0
HA	101	ALA	QB	79111.2
HA	102	SER	HN	6160.0
HA	103	GLU-	HN	5469.0
QB	102	SER	HN	19600.0
102 SER				
HN	102	SER	HA	6285.0
HN	102	SER	HB2	10560.0
HN	103	GLU-	HN	24380.0
HA	102	SER	HB2	5993.0
103 GLU-				
HN	103	GLU-	HA	21050.0
HN	103	GLU-	HB2	34620.0
HN	103	GLU-	HB3	16390.0
HN	103	GLU-	HG2	3696.0
HN	103	GLU-	HG3	3426.0
HA	103	GLU-	HB2	37390.0
HA	103	GLU-	HB3	12350.0
HA	103	GLU-	HG2	1159.0
HA	103	GLU-	HG3	1023.0
HB2	103	GLU-	HG2	21302.4
HB2	103	GLU-	HG3	82790.4
HB3	103	GLU-	HG2	88032.0
HB3	103	GLU-	HG3	45864.0

1D NOE intensities

18 HES				
HA	18	HES	HD1	17700.0
HB2	18	HES	HD1	18100.0
HB3	18	HES	HD1	17600.0
HD1	19	THR	HN	14400.0
HD1	31	ASN	HA	9800.0
67 TYR				
HH	71	PRO	HG3	36700.0
HH	71	PRO	HB2	81400.0
HH	71	PRO	HB3	82200.0

PART II

Irrelevant NOE constraints

HN	PHE	-3	- HA	PHE	-3
HN	PHE	-3	- CG	PHE	-3
HA	PHE	-3	- HB3	PHE	-3
HA	PHE	-3	- CG	PHE	-3
HB2	PHE	-3	- CG	PHE	-3
HB3	PHE	-3	- CG	PHE	-3
CG	PHE	-3	- CZ	PHE	-3
CG	PHE	-3	- HZ	PHE	-3
CZ	PHE	-3	- HZ	PHE	-3
HN	LYS+	-2	- HA	LYS+	-2
HA	LYS+	-2	- HB3	LYS+	-2
HG2	LYS+	-2	- HE2	LYS+	-2
HD2	LYS+	-2	- HE2	LYS+	-2
HN	ALA	-1	- HA	ALA	-1
HA	ALA	-1	- QB	ALA	-1
HN	GLY	1	- HA1	GLY	1
HA	SER	2	- HB2	SER	2
HA	ALA	3	- QB	ALA	3
HN	LYS+	4	- HA	LYS+	4
HA	LYS+	4	- HB2	LYS+	4
HA	LYS+	4	- HN	LYS+	5
HA	LYS+	5	- HB2	LYS+	5
HA	LYS+	5	- HN	GLY	6
HB3	LYS+	5	- HG2	LYS+	5
HN	GLY	6	- HA1	GLY	6
HN	GLY	6	- HA2	GLY	6
HA	ALA	7	- QB	ALA	7
HN	THR	8	- HA	THR	8
HN	THR	8	- QG2	THR	8
HA	THR	8	- HB	THR	8
HB	THR	8	- QG2	THR	8
HA	LEU	9	- HB2	LEU	9
HA	LEU	9	- HN	PHE	10
HB2	LEU	9	- QD1	LEU	9
HB3	LEU	9	- QD2	LEU	9
HA	PHE	10	- HN	LYS+	11
HE1	PHE	10	- HD2	PHE	10
HZ	PHE	10	- HE2	PHE	10
HE2	PHE	10	- HD2	PHE	10
HN	LYS+	11	- HA	LYS+	11
HA	LYS+	11	- HB2	LYS+	11
HA	LYS+	11	- HN	THR	12
HN	THR	12	- HA	THR	12
HA	THR	12	- HN	ARG+	13
HN	ARG+	13	- HA	ARG+	13
HA	ARG+	13	- HB2	ARG+	13
HN	CYSS	14	- HA	CYSS	14
HA	CYSS	14	- HB2	CYSS	14
HA	LEU	15	- HB3	LEU	15
HA	LEU	15	- HN	GLN	16
HG	LEU	15	- QD1	LEU	15
HG	LEU	15	- QD2	LEU	15
HN	GLN	16	- HA	GLN	16
HG2	GLN	16	- HE21	GLN	16
HG2	GLN	16	- HE22	GLN	16
HG3	GLN	16	- HE21	GLN	16
HN	CYSS	17	- HA	CYSS	17
HA	CYSS	17	- HB2	CYSS	17
HA	CYSS	17	- HN	HES	18

HN	HES	18	-	HA	HES	18		HB2	ASN	31	-	HD21	ASN	31
HA	HES	18	-	HB2	HES	18		HB2	ASN	31	-	HD22	ASN	31
HB2	HES	18	-	HD1	HES	18		HB3	ASN	31	-	HD21	ASN	31
HB3	HES	18	-	HD1	HES	18		HB3	ASN	31	-	HD22	ASN	31
HAP71	HES	18	-	HAP72	HES	18		HN	LEU	32	-	HA	LEU	32
HAP71	HES	18	-	HBP73	HES	18		HA	LEU	32	-	HB2	LEU	32
HAP72	HES	18	-	HBP74	HES	18		HB2	LEU	32	-	HG	LEU	32
QM8	HES	18	-	HDM	HES	18		HB2	LEU	32	-	QD1	LEU	32
HDM	HES	18	-	QM1	HES	18		HB2	LEU	32	-	QD2	LEU	32
QM1	HES	18	-	QT2	HES	18		HB3	LEU	32	-	HG	LEU	32
HAM	HES	18	-	QM3	HES	18		HB3	LEU	32	-	QD2	LEU	32
QM3	HES	18	-	HT4A	HES	18		HG	LEU	32	-	QD1	LEU	32
QM3	HES	18	-	QT4	HES	18		HG	LEU	32	-	QD2	LEU	32
HBM	HES	18	-	QM5	HES	18		HN	HIS	33	-	HA	HIS	33
HAP61	HES	18	-	HBP63	HES	18		HA	HIS	33	-	HB2	HIS	33
HN	THR	19	-	HA	THR	19		HB2	HIS	33	-	HD2	HIS	33
HB	THR	19	-	QG2	THR	19		HB3	HIS	33	-	HD2	HIS	33
HA	VAL	20	-	QG1	VAL	20		HN	GLY	34	-	HA2	GLY	34
HA	VAL	20	-	QG2	VAL	20		HN	ILE	35	-	HA	ILE	35
HB	VAL	20	-	QG1	VAL	20		HA	ILE	35	-	QG2	ILE	35
HB	VAL	20	-	QG2	VAL	20		HA	ILE	35	-	HN	PHE	36
HA	GLU-	21	-	HB3	GLU-	21		HB	ILE	35	-	QG2	ILE	35
HA	GLU-	21	-	HG2	GLU-	21		HB	ILE	35	-	HG12	ILE	35
HA	GLU-	21	-	HG3	GLU-	21		HB	ILE	35	-	HG13	ILE	35
HA	LYS+	22	-	HB2	LYS+	22		QG2	ILE	35	-	HG12	ILE	35
HB2	LYS+	22	-	HG2	LYS+	22		HG12	ILE	35	-	QD1	ILE	35
HB3	LYS+	22	-	HG3	LYS+	22		HG13	ILE	35	-	QD1	ILE	35
HG2	LYS+	22	-	HD2	LYS+	22		HN	PHE	36	-	HA	PHE	36
HG2	LYS+	22	-	HE3	LYS+	22		HA	PHE	36	-	CG	PHE	36
HG3	LYS+	22	-	HD2	LYS+	22		HA	PHE	36	-	CZ	PHE	36
HD2	LYS+	22	-	HE3	LYS+	22		HB2	PHE	36	-	CG	PHE	36
HN	GLY	23	-	HA2	GLY	23		HB3	PHE	36	-	CG	PHE	36
HN	GLY	24	-	HA1	GLY	24		CG	PHE	36	-	CZ	PHE	36
HN	GLY	24	-	HA2	GLY	24		CG	PHE	36	-	HZ	PHE	36
HA	PRO	25	-	HB3	PRO	25		CZ	PHE	36	-	HZ	PHE	36
HA	PRO	25	-	HN	HIS	26		HN	GLY	37	-	HA1	GLY	37
HB3	PRO	25	-	HG2	PRO	25		HN	GLY	37	-	HA2	GLY	37
HB3	PRO	25	-	HG3	PRO	25		HA2	GLY	37	-	HN	ARG+	38
HG3	PRO	25	-	HD2	PRO	25		HA	ARG+	38	-	HB3	ARG+	38
HN	HIS	26	-	HA	HIS	26		HA	ARG+	38	-	HG2	ARG+	38
HN	HIS	26	-	HB3	HIS	26		HN	GLN	39	-	HA	GLN	39
HA	HIS	26	-	HB2	HIS	26		HA	GLN	39	-	HB3	GLN	39
HB2	HIS	26	-	HD2	HIS	26		HN	SER	40	-	HA	SER	40
HN	LYS+	27	-	HA	LYS+	27		HN	GLN	42	-	HA	GLN	42
HA	LYS+	27	-	HB2	LYS+	27		HA	GLN	42	-	HB3	GLN	42
HA	LYS+	27	-	HN	VAL	28		HA	GLN	42	-	HN	ALA	43
HB2	LYS+	27	-	HG3	LYS+	27		HB2	GLN	42	-	HG2	GLN	42
HB3	LYS+	27	-	HG3	LYS+	27		HB2	GLN	42	-	HG3	GLN	42
HG3	LYS+	27	-	HD2	LYS+	27		HB3	GLN	42	-	HG2	GLN	42
HD2	LYS+	27	-	HE2	LYS+	27		HB3	GLN	42	-	HG3	GLN	42
HD2	LYS+	27	-	HE3	LYS+	27		HG2	GLN	42	-	HE21	GLN	42
HD3	LYS+	27	-	HE2	LYS+	27		HG3	GLN	42	-	HE21	GLN	42
HN	VAL	28	-	HA	VAL	28		HG3	GLN	42	-	HE22	GLN	42
HA	VAL	28	-	HB	VAL	28		HN	ALA	43	-	HA	ALA	43
HA	VAL	28	-	HN	GLY	29		HA	ALA	43	-	QB	ALA	43
HB	VAL	28	-	QG1	VAL	28		QB	ALA	43	-	HN	GLU-	44
HB	VAL	28	-	QG2	VAL	28		HN	GLU-	44	-	HA	GLU-	44
HN	GLY	29	-	HA1	GLY	29		HA	GLU-	44	-	HG3	GLU-	44
HA	PRO	30	-	HB2	PRO	30		HA	GLU-	44	-	HN	GLY	45
HA	PRO	30	-	HB3	PRO	30		HN	GLY	45	-	HA1	GLY	45
HA	PRO	30	-	HG2	PRO	30		HN	GLY	45	-	HA2	GLY	45
HA	PRO	30	-	HG3	PRO	30		HA1	GLY	45	-	HN	TYR	46
HA	PRO	30	-	HD2	PRO	30		HN	TYR	46	-	HA	TYR	46
HB2	PRO	30	-	HG3	PRO	30		HA	TYR	46	-	HB3	TYR	46
HB3	PRO	30	-	HG3	PRO	30		HB2	TYR	46	-	HD1	TYR	46
HG3	PRO	30	-	HD2	PRO	30		HB3	TYR	46	-	HD1	TYR	46
HN	ASN	31	-	HA	ASN	31		HD1	TYR	46	-	HE1	TYR	46
HA	ASN	31	-	HB3	ASN	31		HD1	TYR	46	-	HE2	TYR	46

HE1	TYR	46	-	HD2	TYR	46	HB3	ASN	63	-	HD22	ASN	63
HE2	TYR	46	-	HD2	TYR	46	HN	MET	64	-	HA	MET	64
HE2	TYR	46	-	HH	TYR	46	HN	SER	65	-	HA	SER	65
HN	SER	47	-	HA	SER	47	HA	SER	65	-	HB3	SER	65
HN	TYR	48	-	HA	TYR	48	HN	GLU-	66	-	HA	GLU-	66
HA	TYR	48	-	HB3	TYR	48	HA	GLU-	66	-	HB3	GLU-	66
HB2	TYR	48	-	HD1	TYR	48	HA	GLU-	66	-	HN	TYR	67
HB2	TYR	48	-	HD2	TYR	48	HB2	GLU-	66	-	HG2	GLU-	66
HB3	TYR	48	-	HD1	TYR	48	HB3	GLU-	66	-	HG2	GLU-	66
HB3	TYR	48	-	HD2	TYR	48	HN	TYR	67	-	HA	TYR	67
HD1	TYR	48	-	HE1	TYR	48	HA	TYR	67	-	HB2	TYR	67
HD1	TYR	48	-	HH	TYR	48	HA	TYR	67	-	HB3	TYR	67
HE1	TYR	48	-	HD2	TYR	48	HD1	TYR	67	-	HE2	TYR	67
HE1	TYR	48	-	HH	TYR	48	HN	LEU	68	-	HA	LEU	68
HE2	TYR	48	-	HH	TYR	48	HB2	LEU	68	-	QD1	LEU	68
HN	THR	49	-	HA	THR	49	HB2	LEU	68	-	QD2	LEU	68
HB	THR	49	-	QG2	THR	49	HB3	LEU	68	-	QD1	LEU	68
HN	ASP-	50	-	HA	ASP-	50	HG	LEU	68	-	QD1	LEU	68
HA	ASP-	50	-	HB2	ASP-	50	HG	LEU	68	-	QD2	LEU	68
HN	ALA	51	-	HA	ALA	51	HN	THR	69	-	HA	THR	69
HA	ALA	51	-	QB	ALA	51	HA	THR	69	-	HB	THR	69
HA	ALA	51	-	HN	ASN	52	HB	THR	69	-	QG2	THR	69
HN	ASN	52	-	HA	ASN	52	HN	ASN	70	-	HA	ASN	70
HA	ASN	52	-	HB3	ASN	52	HA	PRO	71	-	HB2	PRO	71
HB2	ASN	52	-	HD21	ASN	52	HA	PRO	71	-	HB3	PRO	71
HB2	ASN	52	-	HD22	ASN	52	HA	PRO	71	-	HG2	PRO	71
HB3	ASN	52	-	HD21	ASN	52	HA	PRO	71	-	HG3	PRO	71
HB3	ASN	52	-	HD22	ASN	52	HA	PRO	71	-	HD2	PRO	71
HN	ILE	53	-	HA	ILE	53	HA	PRO	71	-	HD3	PRO	71
HA	ILE	53	-	HB	ILE	53	HB2	PRO	71	-	HG2	PRO	71
HA	ILE	53	-	HG13	ILE	53	HB2	PRO	71	-	HG3	PRO	71
HA	ILE	53	-	HN	LYS+	54	HB2	PRO	71	-	HD2	PRO	71
HB	ILE	53	-	QG2	ILE	53	HB2	PRO	71	-	HD3	PRO	71
HB	ILE	53	-	HG12	ILE	53	HB3	PRO	71	-	HG2	PRO	71
HG12	ILE	53	-	QD1	ILE	53	HB3	PRO	71	-	HG3	PRO	71
HG13	ILE	53	-	QD1	ILE	53	HB3	PRO	71	-	HD2	PRO	71
HN	LYS+	55	-	HA	LYS+	55	HB3	PRO	71	-	HD3	PRO	71
HA	LYS+	55	-	HB3	LYS+	55	HG2	PRO	71	-	HD2	PRO	71
HA	LYS+	55	-	HG2	LYS+	55	HG2	PRO	71	-	HD3	PRO	71
HA	LYS+	55	-	HN	ASN	56	HG3	PRO	71	-	HD2	PRO	71
HA	ASN	56	-	HB3	ASN	56	HG3	PRO	71	-	HD3	PRO	71
HB2	ASN	56	-	HD22	ASN	56	HD3	PRO	71	-	HN	TML	72
HB3	ASN	56	-	HD22	ASN	56	HN	TML	72	-	HA	TML	72
HN	VAL	57	-	HA	VAL	57	HA	TML	72	-	HB2	TML	72
HA	VAL	57	-	HB	VAL	57	HA	TML	72	-	HB3	TML	72
HA	VAL	57	-	QG1	VAL	57	HA	TML	72	-	HN	LYS+	73
HA	VAL	57	-	QG2	VAL	57	HG2	TML	72	-	HE2	TML	72
HB	VAL	57	-	QG1	VAL	57	HG3	TML	72	-	HE2	TML	72
HB	VAL	57	-	QG2	VAL	57	HE2	TML	72	-	QQH	TML	72
HN	LEU	58	-	HA	LEU	58	HA	LYS+	73	-	HB3	LYS+	73
HA	LEU	58	-	HB3	LEU	58	HG2	LYS+	73	-	HE2	LYS+	73
HB2	LEU	58	-	QD2	LEU	58	HA	TYR	74	-	HB3	TYR	74
HB2	TRP	59	-	HD1	TRP	59	HA	TYR	74	-	CG	TYR	74
HB2	TRP	59	-	HE3	TRP	59	HB2	TYR	74	-	CG	TYR	74
HD1	TRP	59	-	HE1	TRP	59	HB3	TYR	74	-	CG	TYR	74
HE3	TRP	59	-	HZ3	TRP	59	CG	TYR	74	-	CZ	TYR	74
HE3	TRP	59	-	HH2	TRP	59	HN	ILE	75	-	QG2	ILE	75
HE1	TRP	59	-	HZ2	TRP	59	HA	ILE	75	-	QG2	ILE	75
HZ3	TRP	59	-	HZ2	TRP	59	HA	ILE	75	-	HG12	ILE	75
HZ3	TRP	59	-	HH2	TRP	59	HA	ILE	75	-	QD1	ILE	75
HZ2	TRP	59	-	HH2	TRP	59	HA	PRO	76	-	HB2	PRO	76
HN	GLU-	61	-	HA	GLU-	61	HA	PRO	76	-	HB3	PRO	76
HA	ASN	62	-	HB2	ASN	62	HA	PRO	76	-	HG2	PRO	76
HB2	ASN	62	-	HD21	ASN	62	HA	PRO	76	-	HD2	PRO	76
HB3	ASN	62	-	HD21	ASN	62	HA	PRO	76	-	HN	GLY	77
HB3	ASN	62	-	HD22	ASN	62	HB3	PRO	76	-	HG2	PRO	76
HN	ASN	63	-	HA	ASN	63	HB3	PRO	76	-	HD2	PRO	76
HB3	ASN	63	-	HD21	ASN	63	HG2	PRO	76	-	HD2	PRO	76

HN	GLY	77	- HA1	GLY	77
HA	THR	78	- QG2	THR	78
HB	THR	78	- QG2	THR	78
HB2	LYS+	79	- HG2	LYS+	79
HB2	LYS+	79	- HG3	LYS+	79
HB3	LYS+	79	- HG2	LYS+	79
HB3	LYS+	79	- HG3	LYS+	79
HG3	LYS+	79	- HD2	LYS+	79
HN	ALA	80	- HA	ALA	80
HA	ALA	80	- QB	ALA	80
HN	ALA	81	- HA	ALA	81
HA	ALA	81	- QB	ALA	81
HZ	PHE	82	- QR	PHE	82
HN	GLY	83	- HA2	GLY	83
HB2	LEU	85	- QD1	LEU	85
HG	LEU	85	- QD1	LEU	85
HN	LYS+	86	- HA	LYS+	86
HA	LYS+	86	- HB2	LYS+	86
HN	LYS+	87	- HA	LYS+	87
HN	LYS+	89	- HA	LYS+	89
HA	LYS+	89	- HN	ASP-	90
HN	ASP-	90	- HA	ASP-	90
HA	ASP-	90	- HN	ARG+	91
HN	ARG+	91	- HA	ARG+	91
HA	ARG+	91	- HB3	ARG+	91
HA	ARG+	91	- HN	ASN	92
HB3	ARG+	91	- HG2	ARG+	91
HN	ASN	92	- HA	ASN	92
HA	ASN	92	- HN	ASP-	93
HN	ASP-	93	- HA	ASP-	93
HA	ASP-	93	- HB3	ASP-	93
HA	ASP-	93	- HN	LEU	94
HN	LEU	94	- HA	LEU	94
HA	LEU	94	- HB2	LEU	94
HA	LEU	94	- HB3	LEU	94
HA	LEU	94	- HN	ILE	95
HB2	LEU	94	- QD1	LEU	94
HB2	LEU	94	- QD2	LEU	94
HG	LEU	94	- QD1	LEU	94
HG	LEU	94	- QD2	LEU	94
HN	ILE	95	- HA	ILE	95
HA	ILE	95	- HB	ILE	95
HA	ILE	95	- QD1	ILE	95
HA	ILE	95	- HN	THR	96
HB	ILE	95	- QG2	ILE	95
HB	ILE	95	- HG12	ILE	95
HG12	ILE	95	- QD1	ILE	95
HN	THR	96	- HA	THR	96
HA	THR	96	- HB	THR	96
HA	THR	96	- HN	TYR	97
HB	THR	96	- QG2	THR	96
HA	TYR	97	- HB3	TYR	97
HD1	TYR	97	- HE1	TYR	97
HD1	TYR	97	- HE2	TYR	97
HE1	TYR	97	- HD2	TYR	97
HE2	TYR	97	- HD2	TYR	97
HN	LEU	98	- HA	LEU	98
HA	LEU	98	- HB3	LEU	98
HA	LEU	98	- HN	LYS+	99
HG	LEU	98	- QD1	LEU	98
HG	LEU	98	- QD2	LEU	98
HN	LYS+	99	- HA	LYS+	99
HA	LYS+	99	- HB3	LYS+	99
HA	LYS+	99	- HN	LYS+	100
HB2	LYS+	99	- HG2	LYS+	99
HB2	LYS+	99	- HG3	LYS+	99
HB3	LYS+	99	- HG2	LYS+	99
HB3	LYS+	99	- HG3	LYS+	99
HG2	LYS+	99	- HD2	LYS+	99
HG2	LYS+	99	- HD3	LYS+	99
HG3	LYS+	99	- HD3	LYS+	99
HN	LYS+	100	- HA	LYS+	100
HA	LYS+	100	- HB2	LYS+	100
HA	LYS+	100	- HN	ALA	101
HD2	LYS+	100	- HE2	LYS+	100
HN	ALA	101	- HA	ALA	101
HA	ALA	101	- QB	ALA	101
HN	SER	102	- HA	SER	102
HA	SER	102	- HB2	SER	102
HA	GLU-	103	- HB3	GLU-	103
HA	GLU-	103	- HG2	GLU-	103
HA	GLU-	103	- HG3	GLU-	103
HB2	GLU-	103	- HG2	GLU-	103

PART III

Hydrogen bond constraints

10 PHE			
HN	6 GLY	O	2.40
N	6 GLY	O	3.40
11 LYS+			
HN	7 ALA	O	2.40
N	7 ALA	O	3.40
12 THR			
HN	9 LEU	O	2.40
N	9 LEU	O	3.40
15 LEU			
HN	10 PHE	O	2.40
N	10 PHE	O	3.40
68 LEU			
HN	64 MET	O	2.40
N	64 MET	O	3.40
75 ILE			
HN	71 PRO	O	2.40
N	71 PRO	O	3.40
94 LEU			
HN	90 ASP-	O	2.40
N	90 ASP-	O	3.40
95 ILE			
HN	91 ARG+	O	2.40
N	91 ARG+	O	3.40
97 TYR			
HN	93 ASP-	O	2.40
N	93 ASP-	O	3.40
98 LEU			
HN	94 LEU	O	2.40
N	94 LEU	O	3.40
99 LYS+			
HN	95 ILE	O	2.40
N	95 ILE	O	3.40
101 ALA			
HN	97 TYR	O	2.40
N	97 TYR	O	3.40
102 SER			
HN	98 LEU	O	2.40
N	98 LEU	O	3.40
1 GLY			
HN	96 THR	OG1	2.40
N	96 THR	OG1	3.40
43 ALA			
HN	48 TYR	OH	2.40
N	48 TYR	OH	3.40
41 GLY			

```
HN      18 HES  O2A      2.40
N       18 HES  O2A      3.40
 59 TRP
HE1     18 HES  O2A      2.70
NE1     18 HES  O2A      3.70
```